VALIDITY OF SELF-REPORTED MEDICAL CARE UTILIZATION

ALISON CHRISTINE EDWARDS

VALIDITY OF SELF-REPORTED MEDICAL CARE UTILIZATION

BY

© ALISON CHRISTINE EDWARDS

A thesis submitted to the School of Graduate
Studies in partial fulfilment of the
requirements for the degree of
Master of Science

Division of Community Medicine
Faculty of Medicine
Memorial University of Newfoundland
1992

St. John's Newfoundland

ISBN 0-315-73358-6

ABSTRACT

The objective of this thesis was to study the validity of self-reported medical care utilization. Hospitalization and physician visit data for a twelve month recall period were obtained from both an interview and official records. The self-reported information was collected by a telephone survey applied to all adults over 20 years of age in a probabilistic sample of households in metropolitan St. John's (3,300 subjects, 85% response rate). Verification data were later obtained for 2,994 subjects (91%) from the provincial hospitalization database and health insurance plan.

The utilization data were used to categorize the subjects into those in agreement, the underreporters and the overreporters. A variable denoting level of accuracy was derived. Socio-demographic variables (sex, age and education) and health status variables (self-assessed health status, number of chronic conditions, satisfaction with physical health, and emotional status) were used in a descriptive analysis to compare those in agreement, the under- and overreporters.

Logistic regression was utilized to investigate the probabilities of being in disagreement on utilization and to

compare those who were accurate in self-recall with those who were not.

The analyses showed that most subjects (97.3%) were in agreement on hospitalizations in the recall period whereas 84.1% agreed on physician visits. The observed rates of overreporting were 16.2% for hospitalizations and 9.7% for physicians, and for underreporting, 7.3% for hospitalizations and 10.0% for physicians.

For hospitalizations, 1) subjects in disagreement had, usually, less education and more chronic conditions than those in agreement, 2) underreporters tended to be older than either overreporters or those in agreement, and 3) the less accurate at reporting were generally more likely to report several chronic conditions, fewer years of education, lower emotional and self-assessed health status and were older.

For physician visits, 1) subjects found to be in disagreement were in general male, had better emotional and self-assessed health status, and reported fewer chronic conditions, 2) underreporters tended towards higher self-assessed health status, lower education and fewer chronic conditions than those in agreement, and 3) those more prone to be inaccurate in reporting the number of visits were generally female, had lower self-assessed health status, were less educated and had more chronic conditions.

ACKNOWLEDGEMENTS

This research was supported by Grant No. 6601-1079-46 from the National Health Research Development Program, Department of Health and Welfare, Canada, and by additional funds provided by the Faculty of Medicine, Memorial University of Newfoundland.

Thanks are extended to Dr. Jorge Segovia, Associate Dean of Community Medicine, who was a very supportive supervisor, and to Dr. Roy F. Bartlett, Associate Professor, Department of Mathematics and Statistics, for his advice on the use of logistic regression analysis. Credit is also due to Dr. Robin F. Badgley who suggested, in correspondence to Dr. Segovia, that this topic would be worthy of exploration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

In the early 1950's researchers such as Belloc (1954) and Anderson (reviewed in Anderson, 1991) investigated the validity of self-reported hospital utilization by comparing the information obtained from surveys with that found in hospital records. Since then, with the increase in development and use of computers and databanks, the possibilities of more detailed research have mushroomed. Surveys have become a common research tool in health services and social medicine research and many researchers have been able to analyze very large samples of the population to obtain a realistic picture of the health of any country's population.

Many health interview surveys include questions on medical utilization such as 1) how many visits to hospitals or physicians within a set recall period, 2) how long the hospital visits lasted, 3) details of costs incurred, 4) procedures undertaken, or 5) diagnoses given. All of this autobiographical information can be corroborated by checking the information in the official record sources, be they medical charts, insurance records, hospital records or any other, often computerized, source. The information contained in the official records is taken as being correct and any difference between this response and that collected via the survey instrument is known as the response or recall error.

To perform a validation check is a costly and time-consuming venture and hence it would be of great interest to researchers if those subjects who were more likely to make a response error could be classified on selected criteria, for example socio-demographic or health related characteristics. If this were found to be the case it could be worthwhile considering 1) altering the interviewing process and/or instrument to minimize the errors, 2) validating only those groups shown to make significant errors, or 3) looking at the results from any survey with the knowledge of its inherent deficiencies.

Many studies have endeavoured to identify the characteristics of those who make response errors but most are limited by the lack of a universal medical care plan, and hence records, to ensure that all encounters with the medical care system are included in any analysis of response errors. In Canada, we do have universal health care plans where medical care undertaken by any subject is recorded as an encounter in the subject's home province. Even if visits or hospitalizations occur outside the home province they are, in most cases, still included in the respective database. In Newfoundland, as in other provinces, there is one insurance system which handles all physician visits, either in-patient or out-patient, together with one central database for all hospitalizations. Coverage in these plans is universal,

hence researchers can be assured that virtually all encounters with the health care system will be included in one or the other of these databases.

Health care insurance plans, particularly if universal in coverage, are the ideal source of official records since only one source need be checked for validation of any reported event. Most health insurance plans have been established for the purpose of accounting and remuneration to the medical personnel for work performed, or for remuneration to the patient for costs they have incurred for their health care. Since most people want repayment for costs if they are patients, or payment for duties performed if they are physicians, the insurance plans have a high chance of including all the health care encounters by patients. If universal insurance is in effect all health care encounters will ultimately be paid for by the one insurance plan and therefore all such encounters will be filed in the same place. This enables record checks to be undertaken with the knowledge that all of any patient's medical care events will be listed for comparison purposes. Hence response/record discrepancy studies carried out in Canada, or elsewhere that a universal medical insurance plan exists, have the prospect of being more accurate in comparison than those carried out, for instance, in the USA, where there is no universal medical

plan and hence difficulties of having to obtain validation records from more than one source.

When the official sources have been thoroughly checked for validation of the reported event three scenarios are found: 1) the event is recorded in both the survey and the official records; this then constitutes an 'agreement', 2) the event is recorded in the survey but not in any official source; this is categorized as an 'overreport' or false positive, and 3) the event is found in an official source but was not mentioned in the interview; this is an 'underreport' or false negative. Figure 1.1 shows these three categories and the two sources of information.

Utilization studies undertaken in countries where there are many different official record sources are often of a prospective or retrospective design. In historical prospective studies, an interview is undertaken in which names of providers of health care are obtained, utilization data are then obtained from these providers. Data are not provided for the false negatives (underreports) because events not reported in the interview clearly cannot be verified in the validation check. On the other hand, in retrospective studies, the data are collected from records and then the interview is carried out on the subjects identified; this design does not provide data for the false positives (overreports) as only information on events in the

records are collected, any additional events mentioned in the interview are not further validated in the various medical record sources and hence cannot be classified into overreports or agreements.

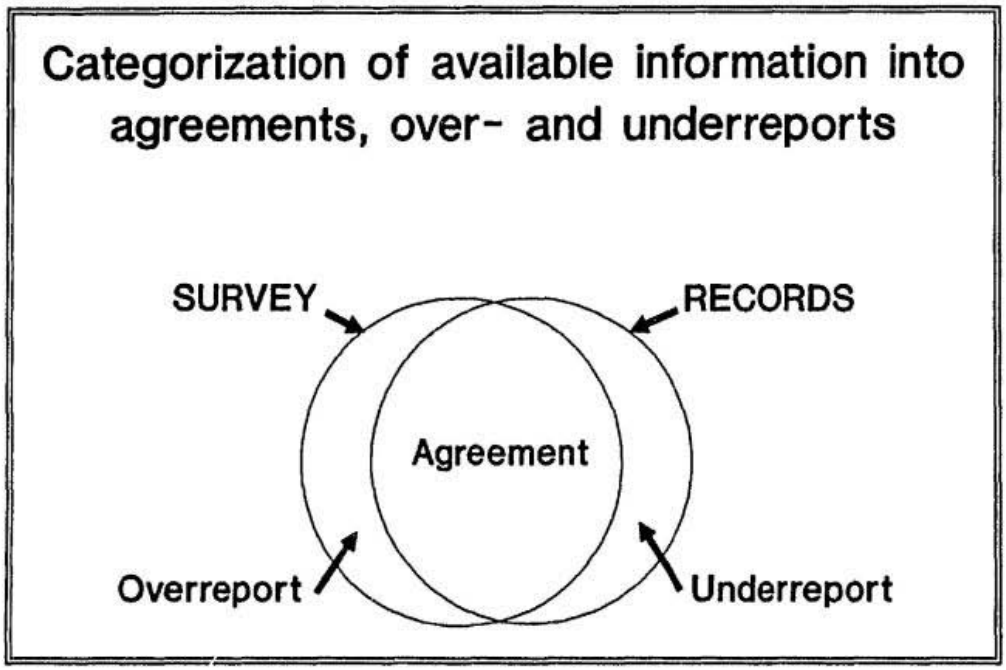


Figure 1.1 Categorization of available information into agreements, over- and underreports

2 LITERATURE REVIEW

2.1 BACKGROUND

2.1.1 GENERAL

Human memory is subject to lapses, obliterations and falsifications and hence recall errors are a factor which must be taken into account particularly when analyzing any survey involving retrospective data. The potential bias can be one of either over- or under-estimation.

Under- and overreports have been analyzed on an individual basis in a number of studies. This avoids the cancelling out of errors between individuals which occurs when a sample or population is considered as a group. Individually, a subject's error is the amount that the interview response is at variance from the validation source; this results in either a negative or positive number. Collectively, for a sample or population, the response errors would be expected to fit a normal curve with a mean of zero. Thus when comparing aggregate data from a survey population, the levels of under- and overreporting will be masked by the cancelling out of the errors, resulting in an overall relative consistency between the averaged survey population data and the validation data. Cleary and Jette (1984) found an average difference of only 0.05 visits between reported physician utilization and actual usage in the year prior to

the interview, whereas in fact approximately 10% of the sample had discrepancies of more than four visits.

One response/record discrepancy study undertaken in Canada was part of an international comparison of medical care. Schach *et al.* (1972) used the government operated hospital services and medical care insurance plans in Saskatchewan to compare questionnaire responses on utilization with the recorded insurance data. Over 99% of the province's population is on the registration files. This guaranteed that a complete and thorough comparison was made between the response and recorded information for the 12 month period under study. Out of the 674 reported and/or recorded hospital admissions, 78% were reported in both the questionnaire and the provincial records. There were 9% underreports found and 13% overreports. These values can be restated as 14% of all questionnaire reported admissions did not have an equivalent entry in the province's records, while 11% of all hospital admissions recorded by the province were not reported by respondents.

Belloc (1954) found 11% overreports and 14% underreports when checking almost 300 hospitalizations in San Jose, California. For physician visits, Andersen *et al.* (1979) found that 35.6% of visits were substantiated in the verification, while 36.6% had overreports and 27.8% underreports. More recently, Jobe *et al.* (1990) found

respondents underreported physician visits by 20%, although they admit this is likely to be an underestimation since unless the provider's name was reported by the subject, the investigators were unable to contact the provider for verification of the report.

Calculating the difference between reported and recorded physician visits in a 12 month period, Mechanic and Newton (1965) showed - in a group of 600 male freshman college students in Wisconsin - that 47% corresponded exactly with medical records. There were 14% who underreported by one visit, 7% by two visits, 4% by three visits, and 12% by four or more. Overreporting was much more unusual; 11% overreported by one visit, and 5% by two or more visits.

Looking at various types of medical care, Yaffe *et al.* (1978) found that in two areas in the USA (Baltimore and Washington County) the services reported with the most accuracy were emergency room visits (94.2% and 96.6% respectively), and inpatient hospital utilization (93.9% and 92.4%). Outpatient clinic visits were reported least accurately (53.5% and 39.1%). Green *et al.* (1979) computed the overall percentage of matched records for various categories of health care and found an 82% match for having seen a general practitioner or other physician in the office, 81% for physician visits to the subject's home (both over a two month recall period), 88% matching on hospital inpatient

stays and 91% for outpatient visits (one year and two month recall respectively).

Coulter *et al.* (1985), in a survey to compare recall of surgical histories, found a 90% concordance rate (265 of 294) in surgical procedures reported in the questionnaire and recorded in the medical records kept on file by the family physician. All surgical procedures in the 35 years preceding the survey were compared. Of the 10% not in agreement, 29 survey responses had no comparable record in the file while 10 were recorded in the files but not given by the respondent. This latter group were all for minor procedures probably carried out in the outpatient departments. For the 29 with no record in the medical files there was suggestive evidence that the deficiency was in the notes rather than in the patient's self-reported history; most dated prior to 1970. The date of surgical procedure was in concordance for 82%; of the other 18%, two thirds were within three years of the correct date.

2.1.2 CLERICAL ERROR AND SUBJECT INTERACTION

If an interviewer, through carelessness, occasionally records an answer incorrectly, an error results. These errors are unlikely to be systematic, and as there may be as many errors in one direction as in the other, over the whole sample they will have little effect (Moser and Kalton, 1972).

With thorough interviewer training and supervision these errors should be minimal.

An interview is an interaction between two people who may affect each other in different ways (Moser and Kalton, 1972). The respondent's attitude towards the interview may influence the effort required to make an accurate report; an accurate report requires the respondent to react in a positive way to the interview and interviewer, and a negative reaction to either may result in inaccurate replies. Those who do not like the interview are less likely to report correctly (Fisher, 1962) and failure of communication may increase the rate of underreporting (Cannell *et al.*, 1977).

2.2 THE MEDICAL CARE VISIT

2.2.1 SIGNIFICANCE AND SEVERITY OF THE EVENT

Accuracy of recall is affected by the severity of the event and the length of time since the event took place (Cherry and Rodgers, 1979). The relevance of the event in question may be such that it is not easily recalled. Physician visits are not particularly salient events in the course of a year, whereas hospital admissions are salient and surgery even more so. In the USA the social class of the patient may influence the place of medical care; persons from lower socioeconomic classes are more likely to use the

clinics while the higher classes use private physicians. Because of this difference, Roghmann and Haggerty (1974) postulated that it could be expected that if someone in the 'lower class' were to see a private physician it would be better remembered than a visit to the clinic, while a telephone consultation by a middle class mother would soon be forgotten especially if followed by a visit to or from the doctor on the following day.

Cannell *et al.* (1965) and Simmons and Bryant (1962) classified illnesses according to seriousness under the headings 'most threatening' 'somewhat threatening' and 'not threatening', and showed that those with 'most threatening' illnesses were more likely to underreport than those with 'not threatening' illnesses (21% versus 10% in Cannell's study). Means *et al.* (1989) noticed that serious events were more likely to be telescoped forward (bringing the date closer to the present) while minor events were put further back in time than they really were. Kulley (1974) and Cannell *et al.* (1977) both found that the involvement of surgery in any hospitalization improved the reporting accuracy, while Schach *et al.* (1972) found that hospitalizations for surgery, child birth and preventive check-up were more reliably reported than those resulting from sickness or injury.

2.2.2 FREQUENCY OF EPISODES

It is reasonable to expect that the more visits one makes to a health facility, the more difficult it would be to remember accurately the precise number. Repeated hospitalizations or medical appointments over an extended period of time may have muddling effects on recall. Persons who use medical services infrequently may have extra difficulty recalling when they did so. Means *et al*. (1989) discussed the theory that where a subject made recurring visits to a health care facility, there would exist a 'generic memory' for the group of events and it would prove difficult to recall individual visits. Mechanic and Newton (1965) and Cleary and Jette (1984) showed that physician underreporting is very substantially related to the number of actual visits made. The former found that in the 3-5 visit group, 10% underreported by at least three visits; while in the 6-7 visit group and the 8 or more visit group, underreporting by three or more visits occurred in 60% and 71% of the cases respectively. The result for overreporting was less clear as only 5% overreported by more than one visit.

It would be expected that using a shorter recall period would assist in the recall of health care encounters. Both Cannell *et al*. (1965) and Cartwright (1963) found very close agreement between the reported and recorded number of

physician consultations in a four week period. Cannell *et al*. included proxy reports, but all of Cartwright's subjects self-reported and she found that underreporting was evident in 5.5% and overreporting in 5.6% of the surveyed population. As 21% of consultations reported at interview were home visits, Cartwright suggests that some of the overreports may be due to non-recording of home visits by the physicians.

Andersen *et al*. (1979) showed that infrequent users of physician services (one visit per year) tend to underreport (i.e. recall no visits) while heavy users (five or more per year) overreport. For overreported physician visits 34% were in error by one visit, 46% by two to five, and 9% by 11 or more. Similar values for underreports were 41% by one visit, 43% by two to five, and 5% by 11 or more. It must be remembered that Andersen admits they would have been unlikely to detect false negatives (i.e. underreports).

For hospitalizations, it was found that the percent of underreporting increased as the number of episodes increased, for those with only one episode 17% underreported, while for those with three or more episodes 24% underreported (NCHS, 1965).

2.2.3 DURATION OF EPISODE

The longer the length of the hospitalization the greater the chance that it will be reported (Fisher, 1962; Simmons

and Bryant, 1962; Schach *et al.*, 1972; Cannell *et al.*, 1977); the shorter the length of hospital stay, the more underreporting will occur. Stays of only one day are much more likely to be underreported than longer stays (Cannell *et al.*, 1965).

Andersen *et al.* (1979) found that 40% of patients reported the correct length of stay, 42% reported longer stays than did the verification data and 18% reported shorter. Of those that overstated their length of stay, 50% erred by only one or two days, while 27% erred by six days or more. Of those who underreported 43% erred by one or two days and 37% by six or more days.

Cannell *et al.* (1965) interviewed a sample of discharged patients for information about hospitalizations during the twelve month period prior to the interview and compared the interview data with hospital record data for the same individuals. They observed a 'heaping effect' where subjects tended to group their statements about timing and duration of hospital stay at five or ten days or multiples of these, and they noted "the tendency towards rounding becomes greater as the number of days increases ... the net effect is for overstated and understated durations to cancel each other out and this leads to average lengths of stay from records and interviews to be in fairly close agreement".

Simmons and Bryant (1962) found that the reported average length of stay was about 2% greater than the average based on records, and for about half of the events the length of stay agreed with that in the record. Yaffe *et al.* (1978) found that the length of stay for hospitalization was reported accurately in surveys in Baltimore (91.5% were accurate) and in Washington County, Maryland (92.4%).

Loftus (1982) and Kulley (1974) noted that more serious illnesses usually require longer hospital stays than those less serious and therefore there may be a confounding effect of severity on length of stay.

2.2.4 LAPSE OF TIME SINCE EVENT

Hermann Ebbinghaus in 1885 (*vide* Baddeley, 1979) was the first to conduct an experiment on memory and he showed the classic relationship between elapsed time and amount forgotten, i.e. the longer the time elapsed the less remembered. The relationship proved to be logarithmic, a result that has been subsequently observed many times under a range of different laboratory conditions (Baddeley, 1979). There are "two classical theories of forgetting, one of which argues simply that memory traces decay spontaneously with time, while the other suggests that forgetting occurs because other material interferes with the retention of the relevant information. Whether or not such interference is a complete

explanation of all forgetting, there is no doubt that interference is a factor" (Baddeley, 1979). Because of this interference one may be able to obtain more reliable and accurate information if one only asks about the most recent episode.

Andersen *et al.* (1979) attempted to reduce the errors due to memory by asking the study participants to check their personal records before reporting the medical experiences. If responses were obtained by proxy, crucial questions concerning hospitalizations were followed up with the patient if the proxy was unable to give detailed information.

Baddeley (1979) conducted an experiment on a group of volunteer subjects who assisted in research in his laboratory, asking them to recall their last visit to the unit. Since he was asking about the last visit only, there should have been no retroactive interference from other visits. They were asked to estimate the date of the last visit and Baddeley showed that the absolute error increased by about twenty days in every 100, with a tendency for subjects to underestimate. In health interview surveys, information requested is not usually restricted to the most recent event, but for hospitalizations (as opposed to physician visits), the most recent event may be the only one that occurred in the recall period since hospitalizations are relatively rare occurrences.

When using a fairly long recall period such as twelve months, there may be inaccurate determination of the period preceding the interview date with subjects reporting events that took place before the recall period (Cannell *et al.*, 1965). It may be difficult to distinguish between a hospitalization in the 11th or 13th month (Zaremba *et al.*, 1985). One important cause of overreporting is the tendency to 'telescope' the events, i.e. move them to a later time period. The chances of reporting a hospitalization decrease as the length of time between the hospitalization and the interview increases (Fisher, 1962). Simmons and Bryant (1962) found there were considerably more hospital admissions reported for the six months immediately prior to the interview than in the seventh to twelfth month, and that for discharges occuring in months 10 and 11 prior to the interview only about half of the episodes were reported. Andersen *et al.* (1979) found that out of 1777 total reported hospital admissions, 163 were found to be outside the survey year.

Cannell *et al.* (1977) compared five studies conducted for the National Center for Health Statistics showed that as the time between event and hospitalization episode increased, the percentage not reported increased from below 5% if the time elapsed was less than two months, up to around 50% nearing the 12 month recall time. The curve for

underreporting increased slowly during the six months following the event, but increased sharply beyond that point. Overreporting of hospital stays also seems to increase as the time lapsed increases (Schach *et al.*, 1972). For visits to physicians, over three-quarters of the reported visits in a two week recall period were accurately dated to within a day (Cannell *et al.*, 1977).

2.3 THE RESPONDENT

2.3.1 SELF RESPONDENT VERSUS PROXY

Responses collected in health surveys are either provided by the respondent (self-reported) or by a proxy, usually a member of the household. It is known that persons report their own experiences more fully than when they give proxy reports (Mechanic and Newton, 1965). Mathiowetz and Groves (1985) proposed two reasons why a self-report may be more accurate than a proxy: 1) the proxy respondent may not know about the event or characteristic in question, and 2) events may not be so salient to the proxy and therefore be omitted or incorrectly remembered. This was only found to affect the reporting of chronic conditions (as opposed to acute conditions) where self-reports were judged 'not serious' more often than proxy reports. Mathiowetz and Groves also proposed two reasons why the proxy may be more

correct than the self-respondent 1) where the proxy is the 'health monitor' in the family or the person who nurses family members or arranges for medical care or pays the medical bills and as such is more aware of the health status and events of the family, and 2) where it may be more acceptable for a proxy to report embarrassing health information about someone else. This latter reasoning was not substantiated in Mathiowetz and Groves' study as no difference was found in the percentages of threatening acute or chronic conditions in self and proxy respondents.

Mathiowetz and Groves carried out a survey to investigate the difference between 1) self-respondents, 2) 'random respondents', where one respondent was chosen at random from all adults residing in the household, and 3) 'knowledgeable adult respondents', where any adult answering the telephone gave information for the household. Single adult households were by definition self-respondents. In the knowledgeable adult case the self-respondents reported more or about the same number of health events for themselves as others, while in the group of random respondents, more health events were reported for other members of the household than for themselves (Mathiowetz and Groves, 1985). They also found that self-reports placed the events at an earlier date than proxies. This is compatible with greater 'forward telescoping' in proxy reports.

Fisher (1962) found that respondent status (self or proxy) was not a significant factor, but others have shown that proxies who report for other family members are more likely to underreport utilization, with poorer reporting for more distant relationships (Cannell *et al.*, 1977). Andersen *et al.* (1979) found that proxy responses were more likely to include adults whose self-assessed health was stated as excellent and those who were said not to worry about their health. Females, older and poorer persons were more likely to be self-respondents.

In a study to compare self-reports and proxies, Enterline and Capt (1959) found that there was no significant difference in the number of chronic conditions reported by the proxies or the subjects themselves, nor in the proportion visiting a physician or hospitalized in the past 12 months. Andersen *et al.* (1979) showed that, overall, self-respondents were only slightly more accurate than proxy respondents, while Kulley (1974) concluded that the small shifts observed in results from proxy to self-response imply that "the use of a family respondent is a reasonable alternative to self-response".

2.3.2 SOCIO-DEMOGRAPHIC STATUS OF RESPONDENT

The better reporters of hospitalizations are those with higher levels of education and/or income and lower ages; sex

seems to show no difference (Cannell *et al.*, 1977). Certain people are more likely than others to underreport information for hospitalizations - persons in low-income families, older persons, nonwhites, and persons with less education (Cannell *et al.*, 1965; NCHS, 1965). From their comparison of five studies Cannell *et al.* (1977) stated that "one cannot generalize that respondents with more education are better at overall reporting than those with less education". Persons 55 years of age or more were more likely to underreport visits to physicians (Cannell *et al.*, 1977). Cleary and Jette (1984) found that younger persons tended to report too many physician visits on average; the average age of those giving accurate reports was 44.7 years, those underreporting by six or more visits was 49.2 years, and those overreporting by six or more visits was 38.4 years. Schach *et al.* (1972) also showed that males, those less than 45 years of age and those with lower socio-economic status tended to overreport.

Family income level is often a better predictor of utilization than either age or education, since income frequently reflects both these variables as well as other motivational components (Cannell *et al.*, 1977). Hospitalizations are better reported as family income increases (Cannell *et al.*, 1965) but the reporting of visits to the physicians shows no such trend.

Andersen *et al.* (1979) found that people in older families, the poor, non-whites and the rural farm population were more likely to overreport hospitalizations later rejected in the verification process. For physician visits, overreporting was more common in the non-whites and the rural farm population.

Fisher (1962) found the influence of family income to be relatively marginal, but agreed with others that black respondents were less likely than others to report hospitalizations.

2.3.3 THE HEALTH OF THE INDIVIDUAL

Cleary and Jette (1984) hypothesized that those most concerned about their health and those having the most disability would overreport their use of physicians; similarly, those who think themselves to be in poor health are also more likely to overreport. Andersen *et al.* (1979) showed that those who do not worry about their health tend to underreport. Zaremba *et al.* (1985) found that the more illnesses a person reports the greater the chance a hospitalization will be reported; the fewer chronic conditions reported the greater the underreporting of hospital episodes ranging from 15% for zero conditions, to 6% if three or more chronic conditions are reported (Cannell *et al.*, 1965).

Manga *et al.* (1987) points out that the unemployed and retired are more likely to experience a hospitalization during the year and they too are more likely to have poorer health status; increasing the number of chronic conditions also increases the likelihood of a hospitalization occurring.

2.3.4 SUPPRESSION OF INFORMATION

Another reason for possible non-recall of events is the suppression of information because of the fear that revealing it would reflect unfavourably on the respondent. This failure of communication could arise from a threat to interviewee self-image, or because of their perceptions of others' attitudes towards them (Fisher, 1962). Cannell *et al.* (1977) postulated that males may underreport illness more than females because by admitting to illness the male subject may threaten his self-image, but this did not prove to be the case, the reporting difference was minimal for both hospitalizations and physician visits.

Loftus (1982) reported that embarrassing hospitalizations (such as diseases of the genital organs) were not reported 21% of the time in comparison to 10% for non-embarrassing stays such as pneumonia. Information may also be suppressed by the respondent because he or she is not motivated to do the interview, because others are present to

hear the responses or because of an underlying reticence to cooperate with strangers (Sudman and Bradburn, 1974).

2.3.5 PERSONAL SIGNIFICANCE OF EVENT

For hospitalizations, memory may not be the only factor involved in whether an event is recalled. The personal significance attributed to the event may affect the chance of it being recalled; the subject may tend to exaggerate or minimize the need for medical care (Zaremba *et al.*, 1985). If he sees himself as a healthy person he is more unlikely to admit to hospitalization, whereas if he is a sick person with an interest in health problems he is likely to give correct information (Fisher, 1962). Cleary and Jette (1984) hypothesized that those who were members of prepaid (universal) medical insurance plans would underreport physician visits because those visits would not be paid for individually. They showed that not being a member of a prepaid plan, and therefore having to pay for each visit, made the subject more prone to overreport. Universal health insurance tends to produce a more casual attitude to utilization in the population than either private health insurance or no insurance. This casual attitude likely translates into inaccurate reporting as Cleary and Jette hypothesized (1984).

2.4 ANALYSIS

Andersen *et al.* (1979) calculated an accuracy score for each variable under analysis. The formula they used was

$$\frac{\text{Social Survey} - \text{Verification}}{\text{Social Survey} + \text{Verification}} = \text{Accuracy Score}$$

A negative result indicated an underreport, while a positive result, an overreport. Accurate reporters were those whose social survey report was within 25% of verification. Using the accuracy score, he found that 88% reported hospital admissions accurately during the survey year, and 82% of those reporting admissions were able to report lengths of stay accurately. The least accurate reporters of hospital admissions and length of stay were the young and old, the poor, non-whites, less educated, those who considered themselves in poor health and who worried about their health, and those who had 11 or more days per admission. Those more likely to overreport include the old, poor, non-whites, less educated and those whose self-assessed health was poor. The underreporters were also the poor, those with less education, and the young.

Those less accurate at reporting physician visits were the poor, non-whites, less educated, those considering themselves in poor health or worried about their health. Because physician visits are not particularly salient events, variables related to memory and salience - levels of use,

whether individuals worried about their health, whether they perceived their health as good or poor - showed the most differences in reporting accuracy for physician visits. Hospitalizations, being more salient events, showed that variables related to personal characteristics (age, education etc.) and motivation, appeared to be important in mis-reporting (Andersen *et al.*, 1979).

Cleary and Jette (1984) used a regression model on the reporting errors for physician visits which showed that advancing age, increased utilization, and being a member of a prepaid plan were related to underreporting, while belief in regular physical checkups, a higher number of chronic conditions and an increase in limiting illness in the past year were associated with overreporting. When the regression is performed on reported and actual utilization separately the results were different; age and number of chronic health problems were significantly associated with utilization for reported but not for actual utilization while, sex and membership of a prepaid plan were significantly associated with utilization for actual but not reported values. Cleary and Jette also looked at the response/record differences in two ways, by using the average reporting error, and by using the average of the absolute difference between reported and actual utilization (i.e., the difference irrespective of its

sign). The second statistic emphasizes the size of the difference and not its direction.

2.5 LIMITATIONS

Inherent problems in validation data sources are that they may not be complete records of an interviewee's medical care utilization. In many countries such as the USA, study participants may have more than one source of validation data, such as various hospitals, family physician clinics, insuring organizations and employers, which makes the job of verifying usage very difficult. Errors may occur in matching an individual's survey and record data and, if no attempt is made to reconcile differences, such mismatches may indicate response errors where none exist (Moser and Kalton, 1972). Inaccuracies in the validation data are an unknown quantity in the spectrum of response errors but, with computerization, should be minimal.

Andersen *et al.* (1979) found that some providers of medical care refused to cooperate with the survey and would not give the validation data necessary for analysis. Many studies asked for details on each utilization event including the names of the insurance company claimed from and the hospital or physician visited. This enabled overreporting to be noticed, but underreports would not have been picked up

unless they were noticed by chance while checking information on other visits. Therefore, reports of non-use (false negatives) are impossible to verify. This problem should not be applicable here in Newfoundland since all visits to physicians (in or outside the province) and hospitals are kept centrally by the Medical Care Commission (MCC) (for physicians) and the Department of Health (for hospitalizations).

Many of the studies in the literature involved relatively small samples in the response/record discrepancy part of the analysis - Coulter *et al.* (1985), 198 people with 386 surgical operations between them; Zaremba *et al.* (1985), 99 were suitably matched; Mechanic and Newton (1965), 600 male freshman college students; Cleary and Jette (1984), 1,026 persons older than 18 years of age with 89% consenting to have their records reviewed; Cartwright (1963), 2,040 individuals; Means *et al.* (1989), only 40 subjects. But other studies had considerably larger sample sizes; for example Andersen *et al.* (1979) surveyed 3,765 families involving 11,619 persons in which 38.3% responded for themselves and the others were proxy responses (children 16 and under did not respond for themselves, and they were 33.9% of the sample). Approximately 10% of the sample refused permission to contact the providers. In the end, verifying data were obtained for over 90% of the hospital admissions

(1,558) and for two-thirds of the physician visits (7,736 visits).

2.6 SUGGESTIONS FOR BETTER REPORTING

In many studies overreporting is not considered as serious a problem as underreporting. Researchers tend to be more concerned about the information subjects fail to report than the reporting of events which did not occur. Many studies have proceeded from the patient record to the interview and as such are unsuitable for evaluating overreporting. Andersen *et al.*, in their study (1979), interviewed the patients first and then attempted to verify reported contacts with the health system. They admitted to their survey being inadequate to evaluate underreporting.

Kars-Marshall *et al.* (1988) compared 11 health interview surveys, and termed the comparison of interview data and that from records as 'criterion validity'. They showed that data on physical performance, acute and chronic diseases, disability and impairment and on use of health facilities showed significant agreement with medical records, physicians statements and/or medical examination. On the other hand, the percentage of agreement on the prevalence of chronic diseases varied widely. For issues in which health planners are interested (e.g. utilization) the health interview

surveys were considered accurate predictors. They concluded that questions on chronic diseases are generally non-valid and that interview surveys are an unsuitable means for measuring chronic diseases in terms of diagnosis. When looking at reliability and recall periods, they state that there is little information on the optimal recall period, but the best data on the use of health services by children were obtained by having the parents keep a calendar for four weeks. Since this was an expensive method, the next best solutions were a two week recall period, followed by one of 12 months.

Cohen *et al*. (1983) examined the estimates for physician visits based on two-week and twelve-month recall periods and found that two-week estimates provided more accurate group estimates, but longer recall periods were more suited if studying individuals, especially since some events may be rare phenomena. They claim that annual estimates, "may be better for classification analysis, i.e., comparing high utilizers with low utilizers of services on a number of other characteristics, but less suited for estimating the actual number of events for individual analysis and modelling".

The investigator can help the interviewee to recall events and to put them in their correct chronological sequence. The accuracy of the subject's memory is likely to be related to the length of the recall period (Cartwright,

1963). One aid to sequencing events is to include in the questions a statement like 'that is, since .....' and stating the date.

Cognitive psychology has, over the past few years, interacted with survey research to provide many theories and helpful suggestions to improve the recall of medical care visits. Jobe and Mingay (1989) in their paper used a 'time line' with 'landmark events' to assist subjects in dating visits to a physician, dentist or others. In addition, a psychological procedure called decomposition was used, where multiple visits could be distinguished individually by suitable probing questions. The two techniques utilized together improved the recall of recurring events to 67% from 32% prior to the procedures being administered.

Means and Loftus (1991) surveyed 143 undergraduates to obtain recall on medical visits and they included a follow-up question asking what method they utilized to obtain the response they gave, 1) recall of individual visits, 2) estimation, or 3) a combination of both. They found that the use of recall of individual events as the method used declined as the number of visits reported increased. Where only one visit was reported, 81% used individual recall compared to 47% when the individual reported three or more visits. This supports the theory that accuracy of medical

care visits decreases as the number of visits increase since more estimation is used for the higher number of visits.

3 METHODS

3.1 DATA COLLECTION

3.1.1 SAMPLE

During seven months spanning 1985/6 the survey "Lifestyle, Health Practices and Utilization of Health Services" was undertaken on a probabilistic sample of households in the metropolitan area of St. John's, Newfoundland, Canada. The sample was obtained using the local telephone directory (1985 version), which was selected as the most suitable and valid sampling frame available at the time. Metropolitan St. John's was defined for the purpose of this study as those residences having St. John's telephone exchanges as listed at the front of the telephone directory. The prefixes for communities outside the area of interest were therefore known and telephone numbers with those prefixes were discarded if sampled, together with any obvious business or institutional numbers within the St. John's area (provincial government, hospitals, university, old age homes, etc.). To obtain the household telephone number, a FORTRAN program was written to produce random numbers from which the page, column and line number in the telephone directory was selected. Telephone coverage in Newfoundland approaches 100% of the population with three to four percent of unlisted numbers. The telephone directory is updated annually. Once a household was selected, all

residents over the age of 20 were invited to participate in the survey; no proxy responses were allowed. The sampling was thus based on a random, single-staged cluster sample with cluster units (households) of unequal size.

The sample size for the survey was estimated in two ways, 1) by considering the marginal distributions for selected variables (sex, age, number of health practices and number of physician visits) to be similar to those found in other surveys and publications, and allowing for a minimal cell size of five, hence obtaining a suggested sample size of 4,320, and 2) by considering the marginal distributions for age (20-44, 45-64 and 65+ years) and sex as noted in the 1981 census data for St. John's, and allowing for analysis to control by the six possible combinations of age and sex, a sample size of 1,836 was reached. Therefore, on the basis of the above calculations, and considering the cost in both time and manpower, a final sample size of 3,000 was selected.

Using the Statistics Canada figure of 2.3 adults per household in St. John's, 1,300 households were needed to attain the required sample size. Since it was expected that half the entries in the directory would not be in our sample frame, and assuming a response rate of 80%, the number of lines selected was increased accordingly.

Contact was attempted at least seven times, over various days of the week and different times of the day, with each

household or individual before that household or individual would be classified as 'no contact'. Refusals were sent a letter explaining the purpose of the study requesting their assistance and a follow-up telephone call was made.

The response rate was 85%, yielding a final sample of 3,300 individuals resident in 1,675 houses. Family houses comprised 83.4% of the total households with single adult households (12.5%) and households of unrelated adults (4.1%) making up the remainder. Table 3.1 shows that the age and sex distributions of the final sample were similar to those stated in the 1986 Census for metropolitan St. John's (Statistics Canada, 1987).

Table 3.1
Comparison between Metropolitan St. John's (1986 Census) and final study sample, sex and age distributions

	1986 Census (%)	Sample (%)
Sex		
male	50.0	45.1
female	50.0	54.9
Age (grouped)		
20 - 44	66.5	66.5
45 - 64	23.9	22.8
65+	9.6	10.7

3.1.2 QUESTIONNAIRE

The questionnaire used in the survey included questions on health practices (sleep, smoking, drinking, exercise, weight and eating breakfast), preventive health, self-

reported height and weight, disability and bed days, self-assessed health and emotional status, chronic conditions, socio-demographic variables and medical care utilization information (Questionnaire included in Appendix). Many questions were obtained from the Canada Health Survey (1981) and the National Survey of Health Practices and Consequences (Wilson and Elinson, 1981) to enable comparisons to be made with other similar research. The questionnaire was pretested on a small sample from the telephone directory. The field work used standard procedures and the questionnaires were edited, coded and entered onto disk with 100% verification. The VAX 8800-VMS (version 4.4) at Memorial University of Newfoundland was utilized for the data storage and analysis. Data analysis utilized SPSS-X (version 3.0), Minitab and BMDP programs. Ethical approval for the survey was given by the Faculty of Medicine at Memorial University. Further information on the sampling protocol and other aspects of the survey is available in Segovia *et al.* (1987) and Veitch (1991).

Within the plans of the study the investigators had included obtaining data relating to physician visits and hospital visits for all subjects who gave verbal consent to do so. Such permission was received from 2,994 (90.7%) subjects.

The question asked of the subjects in the interview to obtain information pertaining to any hospital visits they had made in the recall period was "In the last year (from ____ 1984) have you been a patient in a hospital overnight?" Those who responded "yes" were then asked "How many days did you spend at the hospital?" (Questionnaire included in Appendix). The interviewers were instructed to insert the month of the interview in the first question so that the recall period would be a year; but in effect the period was from the beginning of the month a year previous to the interview, up to the day of the interview itself. Thus the recall period ranged from 365 days (if interviewed on the 1st of the month) to 396 days (if interviewed on the last day of a 31 day month). To be able to divide the cases into underreports, overreports and agreements, the number of days elapsed between the separation date and the interview were calculated. The variance in the recall period unfortunately introduces a questionable area in recall depending on whether the subject took the question as meaning a 365 day period or not. Because of this, some hospitalizations in this questionable period could be misclassified. There were 20 subjects with hospitalizations in this period between the end of the 12 month period and the end of the recall period. If the 12 month period had been taken as the recall period, two underreporters would have been classified as in agreement and

18 subjects in agreement would otherwise have been overreporters.

The variable from the interview used to compare the physician visits to the Medical Care Commission (MCC) records was obtained in response to the question "In the last year, that is from _____ of 1984, did you have a consultation with a doctor?" For those who responded positively, the question was then asked "How many visits did you have in the last year?" As in the hospitalization questions, the interviewers inserted the relevant month in the first question.

Three socio-demographic variables were used to compare the various groups of under- and overreporters. They were sex, age and education. Age was calculated as the interview date minus the date of birth rounded to one decimal place for this analysis. The education questions obtained information on the last grade completed in school as well as any further education undertaken, either trades school or university, and in the case of university the awarding of a degree or not. The questions did not allow the distinction between those currently attending college and those who had completed college, therefore if trades college was marked off it was counted as one extra year of education after secondary school. For university attendance, the awarding of a degree was noted so these subjects could be assumed to have equivalent to 15 years of education, and for those with no

degree currently, the number of years was taken to be thirteen. Thus if only grade school was attended the number of years equals the grade attained (maximum of 11 years in Newfoundland at the time of the survey), for attending trades school or completed diploma courses the education period was taken to be 12 years, for attending university with no degree 13 years of education were assumed, and for a university degree the years were taken as 15. This scaling does not allow for more than one degree or for trades school plus university, but it does give an approximate estimate of the level of education. See Table A1 in the Appendix for frequencies of these socio-demographic variables.

Income level is often used as a socio-demographic variable when investigating medical care utilization, but the question pertaining to income in the survey only gives three broad groupings and was considered not to be a fine enough scale to show any differences. Also, income is not such an important factor when medical care is universal and free to all residents as it is in Newfoundland. The cost of drugs resulting from medical care is not covered by the Province but there are insurance schemes available through many employers which allow coverage for employees and their families.

We have shown previously (Segovia *et al.*, 1989) that there may be three distinct dimensions to health status:

objective (or physical), subjective and emotional. The objective dimension includes 1) the number of chronic conditions, 2) the presence of a disability and 3) the subjects' worry over their health in the past year; the subjective dimension covers 1) satisfaction with physical condition, 2) self-assessed energy level relative to others of their age and 3) self-assessed health status; and the emotional dimension is taken as 1) overall self-assessed happiness level and 2) emotional status as assessed by the Bradburn scale. Self-assessed health status is also associated with the objective dimension to a slightly lesser extent than found in the subjective. Overall, self-assessed health status, alone, is the most useful and informative health status variable (Segovia *et al.*, 1991). For this current analysis only selected health status variables were used.

To see whether the subject's health status had any effect on the accuracy of response, the variables of self-assessed health status (SAHS: excellent, good, fair, poor), number of chronic conditions (CHRCOND: 0, 1, 2, etc.), satisfaction with physical condition (PHYSCOND: very satisfied, satisfied, not too satisfied, not at all satisfied) and emotional status (EMOT: excellent, good, fair, poor) were looked at for the various groups. For all these health status variables it should be noted that the lower the

value the 'better' the response i.e. the better the self-assessed health status, the satisfaction with their physical condition, and their emotional status, and the fewer chronic conditions the subject reported. Table A2 in the Appendix shows frequencies for these variables.

3.1.3 HOSPITALIZATION RECORDS

In the Province of Newfoundland and Labrador, all hospitalizations of registered insured subjects, whether treated within or outside the Province, are compiled by the Province into a provincial database of admissions/separations. Information sources are from, for example, HMRI (Hospital Medical Records Institute) forms and other sources. This information is completed and returned to the Province for all subjects who are currently registered with MCC. The data are made available to researchers under a strict access protocol which, for our study, included proof of consent from the individuals and approval for the project from the Department of Health.

Within two to three years of the initial survey, hospital records were accessed for the fiscal years (April 1 to March 31) 1981/2 up to 1986/7 for the 2,994 (90.7%) who gave us permission. For selected variables, the information from the 9 years of hospitalization data was amalgamated into

one file and sorted by ID number and admission date. This file contained data on the variables shown in Figure 3.1.

ID Number	-	Unique for the survey subject for all utilization components
MCP Number	-	Unique provincial medical care plan number for survey subject
Hospital Number	-	Code used to identify hospital
Chart Number	-	Hospital chart number
Admit Number	-	Hospital admitting number
Admit Date	-	Date of admission
Sex	-	Male/female
Age	-	Age in years
Birth Date	-	Birth date
Separation Type	-	Discharged from hospital alive/dead
Length of Stay	-	Number of days hospitalized
Diagnoses: Primary Secondary Other	-	Diagnoses coded according to ICD-9-CM
Community	-	Standard Statistics Canada geographical community codes for the subject's residence

Figure 3.1 Variables obtained from official records for hospitalizations

From this file the separation date was computed by adding length of stay to the admission date, and reported diagnoses were grouped into the 18 major rubrics of the ICD-9-CM classification (WHO, 1988).

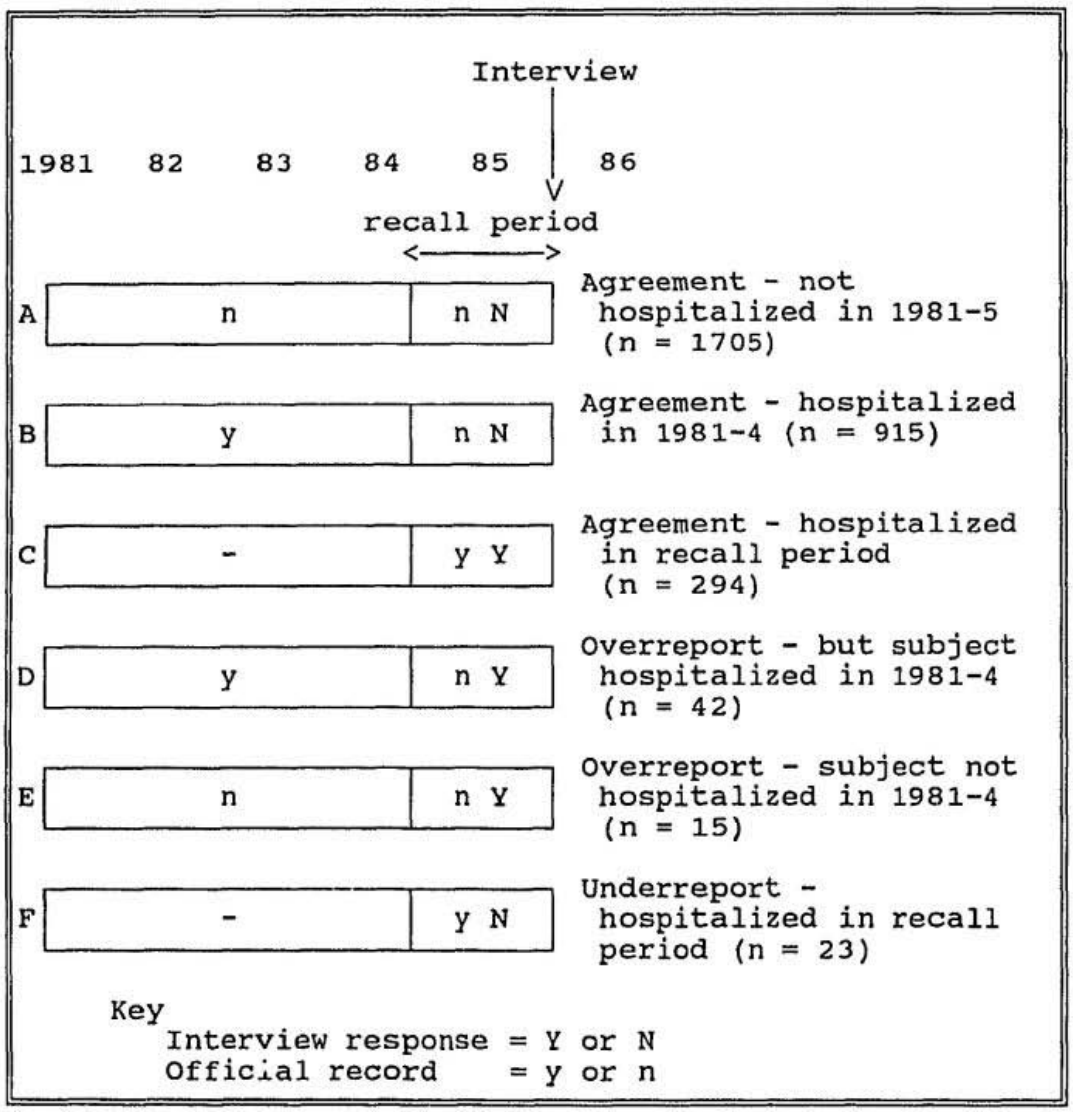


Figure 3.2 Categorization of subjects into groups dependent on reported and recorded visits to hospital

Comparing the existence or not of a hospitalization in the official records with the variable in the questionnaire reporting previous hospitalizations, it was possible to create six categories as shown in Figure 3.2. Combining the first and second group of Figure 3.2 into one group (A+B), and the fourth and fifth into another group (D+E), the variable REPHOSP was developed which divided the 2,994 subjects into four groups as shown in Figure 3.3. Cells 1 (C) and 4 (A+B) together represent the 'agreement' group while cell 2 (D+E) is the 'overreport' group and cell 3 (F) the 'underreport' group.

		Hospitalization record: visit	Hospitalization record: no visit
Interview report	visit	1 (C)	2 (D+E)
	no visit	3 (F)	4 (A+B)

Figure 3.3 Categorization of the variable REPHOSP (letters correspond to those in Figure 3.2)

Taking the hospitalization record as being correct an overall rate for underreporting can be calculated from -

$$100 - \frac{\text{number in agreement on having been hospitalized}}{\text{number hospitalized according to official records}} \times 100$$

and the rate for overreporting from -

$$100 - \frac{\text{number in agreement on having been hospitalized}}{\text{number hospitalized according to interview}} \times 100$$

Another measure of agreement available is Kappa (Harlow and Linet, 1989). It is usually taken that a value of less than 0.20 implies slight agreement, 0.21 to 0.40 fair agreement, 0.41 to 0.60 moderate agreement, 0.61 to 0.80 substantial agreement, and between 0.80 and 1.00 implies almost perfect agreement (Landis and Koch, 1977).

As Figure 3.2 shows, the agreements in cell 4 (no visit/no visit) in Figure 3.3 can be further divided into those who had no hospitalization in the period 1981 to the beginning of the recall period (therefore no hospitalization in their recent memory that could be subject to incorrect recall - A in Figure 3.2) and those who were hospitalized in that period who could have made errors in recall but did not (B in Figure 3.2). The latter are the 'true' agreement group together with those who were hospitalized and who also reported a visit (visit/visit i.e. C in Figure 3.2).

Table A3 in the Appendix supplies information on the 15 overreports who were not hospitalized in the period 1981 to the beginning of the recall period. It should be noted that if any subject only moved to the Province and hence obtained an MCP number in the year previous to the interview, any hospitalization they may have had would not be recorded in

the files. Considering this possibility, the place of birth of the 15 was checked and all were found to have been born in Newfoundland and therefore unlikely to have hospitalizations missed because of not having an MCP number. It is possible that they could have been non-residents of the Province, and therefore not claiming on MCP for the recall period, but we have no way of checking on this possibility. Of these 15 unconfirmed hospital stays, 11 were for one day only and could well have been emergency visits mistakenly reported by the subject to the interviewer. It should be noted that overnight stays in the emergency department are not classified as hospital admissions and therefore there are no HMRI forms completed for these stays.

A summary file containing the subject ID together with the variable REPHOSP was developed and then matched back to the original complete hospitalization file so that all hospitalized subjects could be selected on their value in REPHOSP.

Where a long hospitalization period (of more than 14 days) spanned the day at the start of the recall period, only the portion within the recall period was included to obtain the total number of days hospitalized in the recall period. Hospitalizations of less than 14 days were included in their entirety.

The most recent visit prior to the recall period was used in the analyses for the overreporters. For the underreporters and those in agreement, the visit(s) officially recorded during the recall period were used for the analyses. In both cases where more than one visit was made the lengths of stay were summed. The number of months that the most recent visit preceded the interview was calculated and used to assist in determining if forward telescoping was a possible reason for any overreporting, or lapse of memory over time for the underreporters.

Rather than simply looking at groups of overreporters, underreporters and those in agreement, a different classification was made with respect to accuracy or non-accuracy. It was decided to create the variable ACCHOSP where the 'accurate' group would be those in the agreement group who were either not hospitalized in the recall period or were correct in the number of hospitalization days reported within one day, together with those who overreported their stay by only one day where that stay was in the 13th or 14th month prior to the interview. The 'less accurate' group would consist of all in agreement within two to five days, together with those overreporters in the 13th or 14th month who were in error by two to five days. The 'not very accurate' group would then comprise those subjects in the agreement group who were at discrepancy by more than five

days, the overreporters in the 13th or 14th month at discrepancy by more than five days, and the 'not accurate' group all overreporters who were hospitalized 15 or more months before the interview and all underreporters. These categories are shown in Table 3.2.

Table 3.2
Hospitalizations - classification of subjects into categories of accuracy (ACCHOSP)

	Not in hospital in recall period	In hospital in recall period	Over-reporters	Under-reporters
Accurate	all	diff of ±1 day	diff of ±1 day and visit in 13/14th month	-
Less accurate	-	diff of 2-5 days	diff of 2-5 days and visit in 13/14th month	-
Not very accurate	-	diff of ≥6 days	diff of ≥6 days and visit in 13/14th month	-
Not accurate	-	-	visit in 15+ month	all

The allowance of one day in the criteria for ACCHOSP was partially to offset the differences in the ways lengths of stay were recorded in the survey and official records; subjects were most likely to consider 'days away from home' in recalling the number of days in the hospital, whereas the

official records counted 'nights in hospital' as the length of stay, since the day of discharge was not counted in the tally.

It was possible that subjects may not have recalled hospitalizations for pregnancy when stays in hospital were the subject of recall as they were not caused by an illness or trauma and therefore not in the foreground of the mind while the questions were being asked. Because of this possibility, pregnancy related hospitalizations were excluded from some analyses to see if accuracy of reporting improved with this exclusion.

3.1.4 PHYSICIAN RECORDS

The Medical Care Commission (MCC) in the Province obtains a record for each in- or out-patient visit by a patient to a physician. The records are purely for accounting and repayment purposes, but are available to researchers providing the individual's consent has been obtained and the Commission approves the project. For our survey only those ambulatory visits occurring in either the physician's office, the out-patient clinic, the emergency department or the patient's home were counted. Visits where there are third party payers i.e. insurance companies, or visits for medical examinations for employment are not covered by MCP and would therefore not have been included in

the count obtained. Workers Compensation claims, on the other hand, have been covered by MCP since the early 1980's and were therefore included in the figure obtained; these visits account for a significant proportion of the total visits made to physician offices (personal communication).

The total number of visits to physicians for each study participant for each calendar year was obtained for the years 1985 to 1990. The only information obtained from the Medical Care Commission was the total number of visits in each calendar year for each subject. No dates were obtained for the visits, so a direct comparison of the number of reported visits to physicians with the actual number in the recall period prior to the interview could not be undertaken. To attempt to work around this deficiency, the number of visits over the six years that survey data were collected for, was checked for consistency within each subject; over 75% were consistent within three visits for any year when comparing against their mean number. Since most subjects had a fairly consistent number of visits from year to year, it was assumed that the number of visits in any 12-month period would be comparable to that in a calendar year; therefore for this research, the comparison for validity was done between the number of visits reported and the number in the calendar year 1985 according to the MCC files. The data were edited before this comparison was performed to exclude the recorded zero

visits for years when subjects were known to have been deceased.

It was thought that since the official records gave the number of visits for the calendar year 1985, the interviews conducted in the months July to October might show more differences than those conducted in November to January (recall period closer to the actual calendar year), but this was not evident. Since no clear reasons were evident for excluding any subjects on the basis of month of interview, all subjects were included in the analyses.

As with the hospitalization data, the subjects could be divided into 4 groups (REPMCP) depending on whether any visits were reported during the interview and how many visits were recorded on the official records. In addition to calculating a Kappa value, an overall rate for underreporting was calculated from -

$$100 - \frac{\text{number in agreement on having visited a physician}}{\text{number with visits according to official records}} \times 100$$

and for overreporting from -

$$100 - \frac{\text{number in agreement on having visited a physician}}{\text{number with visits according to interview}} \times 100$$

Accuracy of reporting physician visits was categorized into three groups where the 'accurate' group allowed for an error of one visit more or less than the number reported, the 'less accurate' group, a difference of two to four visits

from that reported, and the 'not accurate' group a difference of five or more visits than that reported. This variable was identified as ACCMCP.

Whereas with hospitalizations it was possible to exclude those visits that were for pregnancy or its related conditions, this was not possible for physician visits as the reason for consultation was not available in our dataset. If those females with pregnancy hospitalizations in the recall period were to be excluded as subjects in the analysis for physician visits, this would have excluded some of the pregnancy visits, but would also exclude other non-pregnancy related visits, and would, on the other hand, not exclude those individuals who attended physicians for pre- or post-natal visits but were not hospitalized for pregnancy in the recall period. For the above reasons, no analyses excluding pregnant subjects were attempted for physician visits.

3.2 ANALYSIS PLAN

Following descriptive analysis and considering the associations and correlations between the discrete independent variables ACCHOSP and ACCMCP and the socio-demographic and health status variables, logistic regression was undertaken on all four dependent variables. The logistic regression (LR) program of BMDP (BMDP, 1988) was used with

selected socio-demographic and health status variables. Resulting regression models were inspected for their goodness of fit by considering the Hosmer-Lemeshow and C.C. Brown chi-square statistics. High p-values for the Hosmer-Lemeshow statistic imply that the predicted values fit the data well, while a high C.C. Brown value signifies that the logistic form is adequate for the data (BMDP, 1988). Where either chi-square p value was less than 0.10 the model was taken as being inadequate for the variables under investigation.

The dependent variables REPHOSP and REPMCP were utilized as dichotomies, comparing 1) those who reported correctly against those who reported incorrectly (underreporters and overreporters) and also comparing 2) under- with overreporters and 3) and 4) both under- and overreporters singly with the agreement group. When a model had been selected the program was run again including only those variables selected in the prior run.

ACCHOSP (four categories) was utilized in two ways, one dichotomy placing those who were 'accurate' or 'less accurate' against those who were 'not very accurate' or 'not accurate', and the other being much more stringent on the accuracy of the report, placing those who were 'accurate' against all others. Because of the large number of subjects who were not hospitalized in any year the logistic regression on ACCHOSP was repeated excluding the subjects who were not

hospitalized in the recall period, since these subjects had no hospitalization to remember correctly and therefore could add nothing to the discriminators for accuracy.

ACCMCP (three categories) was used to compare the accurate group with those who were not accurate. Since there were very few subjects who did not see a physician at least once in any year (as opposed to the case for hospitalizations) no subjects were excluded for any analysis of the MCP data.

Finally, the associations found between the pairs of variables REPHOSP - REPMCP and ACCHOSP - ACCMCP were investigated using Chi-square and Gamma respectively as the measures of association. ACCHOSP and ACCMCP are ordered discrete variables and hence Gamma was a more appropriate measure of association than Chi-square.

4 RESULTS

4.1 HOSPITALIZATIONS

4.1.1 RECALL ERRORS: UNDER- AND OVERREPORTS

Table 4.1 shows the distribution of subjects across the various categories of agreement as shown by the variable REPHOSP. As mentioned in the Methods (section 3.1.2), an overall value for under- and overreporting can be calculated and this is shown in Table 4.2 (for all subjects: underreporting = 100 - (294/317) x 100 = 7.3; and over-reporting = 100 - (294/351) x 100 = 16.2).

Table 4.1
Hospitalizations - comparison of interview report and official record for the recall period prior to the interview (REPHOSP)

Interview report	Hospitalization record	Number	%
Agreement			
No hospitalization	No hospitalization	2620	87.5
Hospitalization	Hospitalization	294	9.8
Overreport			
Hospitalization	No hospitalization	57	1.9
Underreport			
No hospitalization	Hospitalization	23	0.8
Total		2994	100.0

Table 4.2 shows the overall rates of 7.3% underreporting and 16.2% overreporting when taking the hospitalization records as being correct. Females showed a lower rate for

both under- and overreporting than males, and hence the Kappa value indicates a higher rate of agreement. For males and females combined the Kappa value is 0.87 - almost perfect agreement (Landis and Koch, 1977).

Table 4.2
Hospitalizations - overall rates of under- and overreporting together with the Kappa value, by sex

	Male	Female	All
Underreporting (%)	12.1	4.8	7.3
Overreporting (%)	19.7	14.5	16.2
Kappa value	0.82	0.89	0.87

The same length of stay was reported by the subjects and recorded in the verification data for 27.6% of hospitalized subjects; 29.6% reported a shorter length of stay than that recorded and 42.8% reported a longer one. Of those who overreported the length of stay, 61.1% erred by 1-2 days, 20.6% by 3-5 days and 18.3% by six or more days. The equivalent figures for the underreporters were, 66.7% erred by 1-2 days, 11.5% by 3-5 days and 21.8% by six or more days.

The frequencies for the number of days reported by the subjects for hospitalizations showed peaks at seven days, 14 and other multiples of seven; the subjects seemed to often round the stays to the nearest week.

Comparing the percentage who underreported (i.e. reported no hospitalizations) for various numbers of hospital

stays, 7.7% of those with one stay underreported, 4.7% with two stays and 5.9% of those with three or more visits. There was a trend for higher underreporting for subjects with shorter lengths of stay (1-2 days - 7.9%; 8-14 days - 5.0%) but it increased for those who stayed 15 or more days to 8.1%. Those in agreement showed the opposite trend to the underreporters, 71.1% in agreement for those with 1-2 days, rising to 87.5% for those with 8-14 days but declining to 81.1% for 15 or more days.

For hospitalizations officially recorded one to three months before the interview 2.5% of subjects underreported, this percentage increased to 3.9% in months 4-6, 5.6% in months 7-9 and 15.2% for months ten or more.

The first line of 'agreement' in Table 4.1 can be further subdivided into those who were never hospitalized 1981-4 (n=1,705) and those who were hospitalized 1981-4 (n=915) (see Figure 3.2 for clarification). As mentioned in the Methods, this latter group could be considered as the 'true' group of not hospitalized agreements as they correctly remembered that their hospitalization was prior to the recall period. This group of 915 can be compared to the 57 overreporters who incorrectly remembered the placement of their hospitalization. For the group in agreement for having being hospitalized in the recall period, the difference in

the number of days spent in hospital ranges from an underestimate of 90 days to an overestimate of 64 days.

First, looking at the underreporters, it was found that the 23 subjects had a total of 73 hospitalizations between April 1981 and the interview date. Considering only those episodes in the recall period prior to the interview, there were 27 forgotten episodes for the same 23 subjects. One subject had three visits while two had two visits each.

As underreporting can be due to sensitivity of diagnosis or treatment (Loftus, 1982), the primary diagnosis for each visit was listed. Sensitive diagnoses found included two cases of alcohol dependency syndrome, one subject with a vasoplasty and two subjects with genitourinary diagnoses. On the other hand, it was surprising that some hospitalizations were forgotten, for example heart failure (in one case), and the birth of a child (in two subjects). As stated previously, hospitalizations for delivery could have been omitted at recall because the subjects were focusing on illness and trauma. Some underreports were of fairly long duration (minimum one day and maximum 28 days). The two longest stays were both for alcohol dependency syndrome, possibly intentionally omitted. The difference in days between the recorded hospitalization and the interview date was calculated. Table A4 in the Appendix lists the underreports.

For the 57 overreports, the hospital data file was inspected to find any hospitalizations between April 1981 and the interview date. A total of 78 hospitalizations were found for 42 subjects. The difference in days between the most recent hospitalization and the interview ranged from 379 to 1,471 days. Detailed information on these 42 cases is shown in Table A5 in the Appendix. Subtracting 42 from the 57 overreporters left 15 subjects who said they were hospitalized in the previous year; but according to the hospital files, had never been hospitalized since 1981 (a four to five year period).

Table A5 in the Appendix includes the number of days reported by the individual together with the number of days recorded in the hospitalization file. It is interesting to note the proximity for most subjects. These subjects seem to have remembered the hospitalization accurately but misplaced it in time.

Many of the diagnoses in these overreporters could be classified as recurrent and/or chronic diagnoses (Means and Loftus, 1991); it would be expected that these events may be more prone to forward telescoping and less complete recall than 'one-time' non-recurring and/or acute events. For example, within those who were hospitalized in the 13th or 14th month prior to the interview, six cases could be described as chronic by the NHIS criteria (Brewer *et al.*,

1989, Means and Loftus, 1991) (subjects 244101, 818203, 1019101, 219101, 872601, 1245102). Pregnancy related diagnoses account for two of the remaining three (4101, 1510101); these could be linked mentally to the presence of the young baby in the household and therefore prone to forward telescoping.

4.1.2 SOCIO-DEMOGRAPHIC VARIABLES

Table 4.3
Hospitalizations - underreports, socio-demographic variables

	All under-reporters	Excluding alcoholism and vasoplasty	Excluding all sensitive diagnoses
SEX			
male (%)	56.5	50.0	50.0
female (%)	43.5	50.0	50.0
AGE (years)			
$\bar{x}$	54.4	55.7	54.3
SD	19.9	20.8	21.5
SE	4.1	4.7	5.1
EDUCATION (years)			
$\bar{x}$	9.5	9.4	9.7
SD	3.4	3.4	3.4
SE	0.7	0.8	0.8
(n)	(23)	(20)	(18)

Table 4.3 looks descriptively at the 3 variables SEX, AGE and EDUCATION for the underreporters. Considering all those who underreported, there were slightly more males than

females; the mean age was 54.4 years while the mean education was 9.5 years (range three to 15 years).

When looking at those who overreported hospitalizations, Table 4.4 shows that there were very few differences between those who were hospitalized sometime in the past and those who were never in the hospital since April 1981. There were slightly more females in the group who were actually hospitalized in the previous four years compared to those never hospitalized in those four years, but the difference was not significant ($p>0.10$). EDUCATION and AGE were very similar with no significant differences ($p>0.10$).

Table 4.4
Hospitalizations - overreports, socio-demographic variables

	All over-reporters	Those in hospital since 1981	Those never in hospital since 1981
SEX			
male (%)	40.4	38.1	46.7
female (%)	59.6	61.9	53.3
AGE (years)			
$\bar{x}$	41.8	42.2	40.6
SD	18.5	19.8	14.8
SE	2.5	3.1	3.8
EDUCATION (years)			
$\bar{x}$	10.2	10.0	10.9
SD	2.8	2.7	2.9
SE	0.4	0.4	0.8
(n)	(57)	(42)	(15)

For the 42 subjects who were hospitalized in the period of more than one year prior to the interview, the number of

months elapsed since the separation from hospital to the interview date was calculated (maximum 49 months). Subjects were then divided into those who were hospitalized in the two months before the beginning of the recall period and those at more discrepancy.

Table 4.5
Hospitalizations - overreports where subjects were hospitalized prior to recall period before interview, socio-demographic variables

	All over-reporters	Those in 13th or 14th month	Those in 15th or greater month
SEX			
male (%)	38.1	22.2	42.4
female (%)	61.9	77.8	57.6
AGE (years)			
$\bar{x}$	42.2	47.5	40.8
SD	19.8	22.3	19.2
SE	3.1	7.4	3.3
EDUCATION (years)			
$\bar{x}$	10.0	10.4	10.2
SD	2.7	2.3	2.7
SE	0.4	0.5	0.5
(n)	(42)	(9)	(33)

Table 4.5 shows statistics for all the overreporters together and the group categorized by length of time elapsed. The mean age was found to be 42.2 years. The youngest mean age was found in the subgroup who were inpatients in the period of more than 15 months before the interview. The EDUCATION was comparable for the various subgroups. There

were more females than males to be found in all groups of overreporters. It should be noted that there were only nine subjects in one of the comparison groups so any conclusions made from this table should be viewed with caution.

Comparing the overreporters to the underreporters (left hand columns in Tables 4.3 and 4.4) there were significant differences found in AGE for the two groups between all the underreporters and all the overreporters. The overreporters were younger than the underreporters ($t(38) = 2.61$, $p<0.05$). There was no difference found for either SEX or EDUCATION level ($p>0.10$).

As stated earlier, the subjects who agreed on hospitalization in the recall period could be divided into two groups. Table 4.6 shows the socio-demographic variables for each of the two groups and for them combined. It can be seen that there were significant differences in the SEX distribution, mean AGE and EDUCATION level for the two groups (SEX: $X^2(1) = 133.67$, $p<0.0001$; AGE: $t(1693) = 3.93$, $p<0.001$; EDUCATION: $t(1845) = 2.97$, $p<0.01$). The differences between those hospitalized and those not hospitalized are plausible and as expected since 1) more females than males are hospitalized in any year, 2) the hospitalized are older (older subjects are more likely to be hospitalized than younger subjects) and 3) those hospitalized have lower levels of education.

Table 4.6
Hospitalizations - agreements on no hospitalization in recall period, socio-demographic variables

	Not in hospital in the recall period	Not in hospital 1981 to recall period	In hospital during 1981 to recall period
SEX			
male (%)	46.5	54.8	31.1
female (%)	53.5	45.2	68.9
AGE (years)			
$\bar{x}$	40.5	39.6	42.1
SD	15.3	14.6	16.4
SE	0.3	0.4	0.5
(n)	(2620)	(1705)	(915)
EDUCATION (years)			
$\bar{x}$	11.4	11.5	11.2
SD	2.5	2.5	2.5
SE	0.05	0.1	0.1
(n)	(2611)	(1702)	(909)

Table 4.7 shows the descriptive statistics for the survey population who were not hospitalized and also stated they were not in hospital and for those subjects where there was agreement between the reported and recorded hospitalization events (for those who were hospitalized). There was a higher percentage of females in the hospitalized group than in the non-hospitalized group. (This was not unexpected since a higher proportion of females are hospitalized due to pregnancy related episodes). Significant differences were found between the two groups for AGE, SEX and EDUCATION (AGE: $t(346) = 2.23$, $p<0.05$; SEX: $X^2(1) =$

22.50, $p<0.0001$; EDUCATION: $t(350) = 3.17$, $p<0.01$). The hospitalized were older, had less education and were more likely to be female.

Table 4.7
Hospitalizations - agreements - hospitalized versus the non-hospitalized, socio-demographic variables

	Agreement on no hospitalization	Agreement on hospitalization
SEX		
male (%)	46.5	32.0
female (%)	53.5	68.0
AGE (years)		
$\bar{x}$	40.5	42.8
SD	15.3	17.4
SE	0.3	1.0
(n)	(2620)	(294)
EDUCATION (years)		
$\bar{x}$	11.4	10.9
SD	2.5	2.7
SE	0.05	0.2
(n)	(2611)	(294)

Table 4.8 compares the two 'real' groups of agreements, i.e. those who remembered correctly that they were in hospital and those who remembered correctly that their hospital visit was before the recall period. There was no difference in either the SEX or the AGE distributions but the EDUCATION was slightly different with those agreeing on a hospitalization in the recall period having slightly fewer years of education ($t(463) = 1.83$, $p<0.10$).

Table 4.8
Hospitalizations - agreements - hospitalized versus the non-hospitalized (but in hospital between 1981 and the recall period), socio-demographic variables

	Agreement on no hospitalization in recall period	Agreement on hospitalization in recall period
SEX		
male (%)	31.1	32.0
female (%)	68.9	68.0
AGE (years)		
$\bar{x}$	42.1	42.8
SD	16.4	17.4
SE	0.5	1.0
(n)	(915)	(294)
EDUCATION (years)		
$\bar{x}$	11.2	10.9
SD	2.5	2.7
SE	0.1	0.2
(n)	(909)	(294)

Within the agreement group, the accuracy of reporting length of stay (accurate or not according to the variable ACCHOSP) could be looked at. Table 4.9 shows that when comparing those who were accurate about the number of days reported (difference of 0 or ±1) to those who showed more discrepancy, there was a highly significant difference in both AGE and EDUCATION level (AGE: $t(194) = 5.29$, $p<0.0001$; EDUCATION: $t(180) = 5.35$, $p<0.0001$). SEX was almost significant ($X^2(1) = 3.76$, $p<0.10$). The more accurate group was younger, had attained a higher education level and was more likely to be female.

Table 4.9
Hospitalizations - agreements - accuracy of reporting length of stay, socio-demographic variables

	All agreements	Difference of 1 day or less	Difference of more than 1 day
SEX			
male (%)	32.0	27.8	38.6
female (%)	68.0	72.2	61.4
AGE (years)			
$\bar{x}$	42.8	38.5	49.6
SD	17.4	14.6	19.2
SE	1.0	1.1	1.8
(n)	(294)	(180)	(114)
EDUCATION (years)			
$\bar{x}$	10.9	11.6	9.8
SD	2.7	2.1	3.1
SE	0.2	0.2	0.3
(n)	(294)	(180)	(114)

Comparing the underreporters with the agreements (left hand columns in Table 4.3 and Table 4.9), the underreporters were older ($t(25) = 2.71$, $p<0.05$), more likely to be male ($X^2(1) = 5.75$, $p<0.05$) and had marginally less years of education ($t(25) = 1.87$, $p<0.10$) than those who were in agreement. There were no significant differences between the overreporters and those in agreement.

4.1.3 HEALTH STATUS VARIABLES

Health status variables showed very little difference for the various groupings of under- and overreporters (Tables

A6, A7 and A8 included in Appendix for reference). Tables 4.10 and 4.11, on the other hand, do show differences.

Table 4.10
Hospitalizations - agreements - hospitalized versus the non-hospitalized, health status variables

	Agreement on no hospitalization	Agreement on hospitalization
SAHS		
$\bar{x}$	1.9	2.1
SD	0.7	0.8
SE	0.01	0.04
(n)	(2620)	(294)
CHRCOND		
$\bar{x}$	1.0	1.5
SD	1.2	1.5
SE	0.02	0.09
(n)	(2620)	(294)
PHYSCOND		
$\bar{x}$	2.0	2.1
SD	0.6	0.7
SE	0.01	0.04
(n)	(2619)	(294)
EMOT		
$\bar{x}$	1.6	1.8
SD	0.7	0.8
SE	0.01	0.05
(n)	(2612)	(294)

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

Significant differences were found (Table 4.10) for all three variables when comparing the 2,620 not hospitalized to

the 294 who were hospitalized (SAHS: $t(349) = 4.31$, $p<0.0001$; CHRCOND: $t(334) = 6.22$, $p<0.0001$; PHYSCOND: $t(346) = 2.24$, $p<0.05$; EMOT: $t(341) = 4.07$, $p<0.001$). Self-assessed health status (SAHS) was better in the non-hospitalized, they had fewer chronic conditions, they were more satisfied with their physical condition and had a better emotional score than the hospitalized.

Table 4.11
Hospitalizations - agreements - accuracy of reporting length of stay, health status variables

	All agreements	Difference of 1 day or less	Difference of more than 1 day
SAHS			
$\bar{x}$	2.1	1.9	2.3
SD	0.8	0.7	0.8
SE	0.04	0.1	0.1
CHRCOND			
$\bar{x}$	1.5	1.3	2.0
SD	1.5	1.4	1.6
SE	0.1	0.1	0.1
PHYSCOND			
$\bar{x}$	2.1	2.1	2.1
SD	0.7	0.6	0.8
SE	0.04	0.05	0.1
EMOT			
$\bar{x}$	1.8	1.7	2.0
SD	0.8	0.7	0.9
SE	0.05	0.1	0.1
(n)	(294)	(180)	(114)

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

For Table 4.11, significant differences were found for SAHS ($t(230) = 4.45$, $p<0.0001$), CHRCOND ($t(214) = 4.03$, $p<0.001$) and EMOT ($t(195) = 2.58$, $p<0.05$) but not for PHYSCOND. Those who were accurate to within one day had better self-assessed health status, had fewer chronic conditions and had a better emotional score than those who were less accurate.

Table 4.12 compares the health status variables for the agreement group who were not in hospital during the recall period divided into those who were hospitalized during 1981-4 and those who were not.

As would be expected, those who were in hospital in the three years prior to the recall period had worse SAHS ($t(1761) = 4.59$, $p<0.0001$), more chronic conditions ($t(1547) = 7.08$, $p<0.0001$), and were less satisfied with their physical condition ($t(1796) = 2.08$, $p<0.05$) than those who were hospitalized during that period; there was no difference in their emotional score.

Table 4.12
Hospitalizations - agreements on no hospitalization in recall period, health status variables

	Not in hospital in the recall period	Not in hospital 1981 to recall period	In hospital during 1981 to recall period
SAHS			
$\bar{x}$	1.9	1.9	2.0
SD	0.7	0.7	0.7
SE	0.01	0.02	0.02
(n)	(2620)	(1705)	(915)
CHRCOND			
$\bar{x}$	1.0	0.8	1.2
SD	1.2	1.1	1.3
SE	0.02	0.03	0.04
(n)	(2620)	(1705)	(915)
PHYSCOND			
$\bar{x}$	2.0	2.0	2.0
SD	0.6	0.6	0.7
SE	0.01	0.02	0.02
(n)	(2619)	(1704)	(915)
EMOT			
$\bar{x}$	1.6	1.6	1.6
SD	0.7	0.6	0.7
SE	0.01	0.02	0.02
(n)	(2612)	(1699)	(913)

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

Table 4.13
Hospitalizations - agreements - hospitalized versus the non-hospitalized (but in hospital between 1981 and the recall period), health status variables

	Agreement on no hospitalization in recall period	Agreement on hospitalization in recall period
SAHS		
$\bar{x}$	2.0	2.1
SD	0.7	0.8
SE	0.02	0.04
(n)	(915)	(294)
CHRCOND		
$\bar{x}$	1.2	1.5
SD	1.3	1.5
SE	0.04	0.09
(n)	(915)	(294)
PHYSCOND		
$\bar{x}$	2.0	2.1
SD	0.7	0.7
SE	0.02	0.04
(n)	(915)	(294)
EMOT		
$\bar{x}$	1.6	1.8
SD	0.7	0.8
SE	0.02	0.05
(n)	(913)	(294)

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

Table 4.13 shows the group who agreed on hospitalization in the recall period to have lower SAHS ($t(473) = 2.26$,

$p<0.05$), more chronic conditions ($t(451) = 3.34$, $p<0.001$) and lower emotional ratings ($t(455) = 3.95$, $p<0.001$) than those who were hospitalized between 1981 and the recall period. The two groups were comparable in their level of satisfaction with their physical condition.

Comparing the underreporters with the agreements (Tables A6 and 4.11) the underreporters had marginally more chronic conditions than those who were in agreement ($t(25) = 1.89$, $p< 0.10$). There were no significant differences for any health status variable looked at between the overreporters and those in agreement or between the under- and overreporters.

4.1.4 EXPLANATORY ANALYSIS

Table 4.14
Hospitalizations - summary variable accuracy (ACCHOSP) by sex

ACCHOSP	Male		Female	
	n	%	n	%
Accurate	1269	94.1	1531	93.1
Less accurate	27	2.0	49	3.0
Not very accurate	19	1.4	18	1.7
Not accurate	34	2.5	37	2.2

Together with the variable REPHOSP which has been analyzed in the above section, another summary variable ACCHOSP was developed as described in the Methods (p47). This variable is an ordered discrete variable and its distribution by sex is shown in Table 4.14.

Before commencing on logistic regression analysis to investigate the effect the independent variables had on the dependent variables (REPHOSP and ACCHOSP), the independent variables were studied to see their associations (Gamma) and inter-correlations to help decide which to include in the subsequent analysis.

Gamma and Spearman correlation values were obtained between the three socio-demographic variables, the four health status variables and the dependent variable ACCHOSP. The other dependent variable REPHOSP is a non-orderable discrete variable and was therefore not suited to correlation nor Gamma analyses.

Table 4.15 lists the values of Gamma for the variables against ACCHOSP. For this table grouped versions of the variable AGE (20-44,45-64, 65+) and EDUCATION (<high school, high school, trades school or diploma, university - no degree, university - with degree) were utilized. The differences by SEX were noticeable especially for the variables AGE, EDUCATION and CHRCOND. It was decided to see if the inclusion of pregnancy related hospitalizations was responsible for the differences found between the sexes so the Gammas were obtained with these cases excluded (n = 92) and it was found that the differences were then reduced and in some cases eliminated (Table 4.15). Excluding these cases from Table 4.14 increased the number in the 'accurate' group

to 1,462 (94.1%) and reduced the 'less accurate' to 33 (2.1%), the 'not very accurate' to 27 (1.7%) and the 'not accurate' to 31 (2.0%).

Table 4.15
Hospitalizations - associations (Gamma) between each of the socio-demographic and health status variables and accuracy (ACCHOSP)

	All n=2994	Male n=1349	Female n=1645	All excl preg n=2902	Female excl preg n=1553
SEX	0.08	-	-	-0.01	-
AGE (grouped)	0.39	0.49	0.30	0.47	0.45
EDUCATION (grouped)	-0.32	-0.40	-0.26	-0.36	-0.32
SAHS	0.37	0.38	0.37	0.41	0.45
CHRCOND	0.42	0.47	0.39	0.47	0.49
PHYSCOND	0.07	0.03	0.11	0.07	0.12
EMOT	0.30	0.30	0.30	0.32	0.33

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

Table 4.16 shows that of the socio-demographic variables under consideration, EDUCATION and AGE were highly correlated to each other. For the health status variables, SAHS and CHRCOND were correlated with AGE and EDUCATION and CHRCOND, and PHYSCOND and EMOT were correlated with SAHS. Because of the large sample size (n = 2,994) most variables correlated

at some level with the others but for the purpose of selecting the more significant correlations those where the coefficient has at least an absolute value of 0.20 are mentioned. Separate runs by SEX produced similar values, but when all pregnancy related hospitalizations were excluded, most variables showed a stronger correlation to ACCHOSP than when the pregnancy episodes were included (Table 4.17).

Since omitting the pregnancy related episodes did not make much improvement to the correlations, these episodes were not excluded in the logistic regression analysis. The variable SEX was included as an independent variable in the models and only where it was entered into the model was the logistic regression then run separately by sex.

The selected socio-demographic and health status variables for use in the logistic regression analysis are shown in Figure 4.1. Three socio-demographic (SEX, AGE and EDUCATION) and three health status (SAHS, CHRCOND and EMOT) variables were included as independent variables in all runs. PHYSCOND was omitted as it showed the least association and correlation with other variables. AGE and EDUCATION were included as continuous variables to maximize their influence on the dependent variable.

Table 4.16
Hospitalizations - correlations between variables under consideration (coefficient and significance level)

	AGE (yrs)	EDUC (yrs)	SEX	SAHS	CHRCOND	PHYSCOND	EMOT
EDUC (yrs)	-.3501 .000						
SEX	.0360 .025	-.0459 .006					
SAHS	.1160 .000	-.2520 .000	.0011 .475				
CHRCOND	.2870 .000	-.1634 .000	.1384 .000	.3260 .000			
PHYSCOND	-.0724 .000	.0662 .000	-.0169 .178	.3548 .000	.1763 .000		
EMOT	-.1107 .000	-.0526 .002	.0383 .018	.2183 .000	.1503 .000	.2102 .000	
ACCHOSP	.0871 .000	-.1226 .000	.0192 .147	.1154 .000	.1527 .000	.0207 .129	.0910 .000

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status
ACCHOSP: accuracy of reporting hospitalizations

Table 4.17
Hospitalizations - correlations between variables under consideration, excludes pregnancy related hospitalizations (coefficient and significance level)

	AGE (yrs)	EDUC (yrs)	SEX	SAHS	CHRCOND	PHYSCOND	EMOT
EDUC (yrs)	-.3575 .000						
SEX	.0597 .001	-.0533 .002					
SAHS	.1158 .000	-.2513 .000	.0091 .312				
CHRCOND	.2929 .000	-.1636 .000	.1507 .000	.3279 .000			
PHYSCOND	-.0735 .000	.0602 .001	-.0184 .161	.3601 .000	.1795 .000		
EMOT	-.1101 .000	-.0546 .002	.0350 .030	.2240 .000	.1512 .000	.2186 .000	
ACCHOSP	.1153 .000	-.1321 .000	.0020 .456	.1242 .000	.1645 .000	.0199 .142	.0926 .000

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status
ACCHOSP: accuracy of reporting hospitalizations

Independent variables		Design variables
SEX	male	0
	female	1
AGE (yrs)	range 20 - 87	
EDUC - education (yrs)	range 0 - 15	
SAHS - self-assessed health status		
	excellent	0 0 0
	good	1 0 0
	fair	0 1 0
	poor	0 0 1
CHRCOND - number of chronic conditions	range 0 - 8 (only one subject reported 8 conditions, so Figures only show 0 - 7)	
EMOT - emotional status		
	excellent	0 0 0
	good	1 0 0
	fair	0 1 0
	poor	0 0 1

Figure 4.1 Independent variables used in logistic regression analyses

Figure 4.2 lists the dependent variables with their dichotomies used for the logistic regressions analyses on REPHOSP and ACCHOSP.

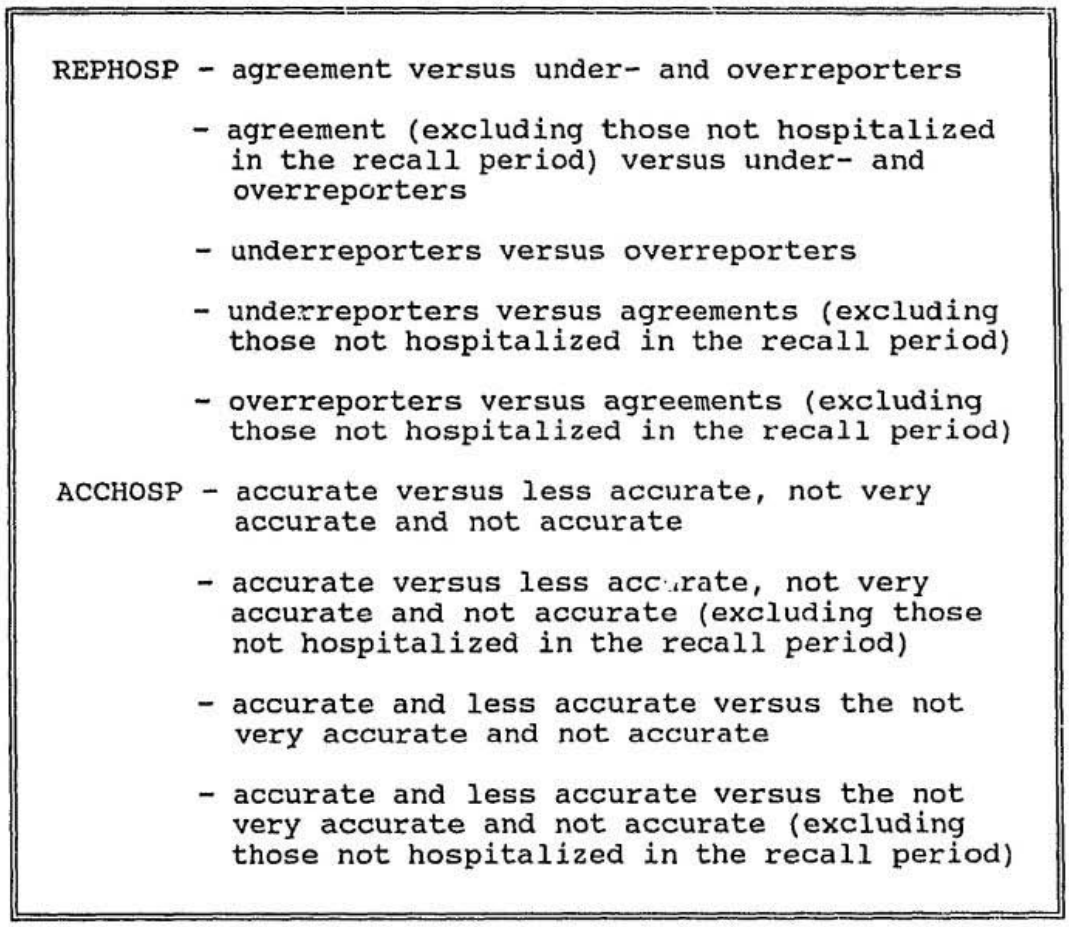


Figure 4.2 Hospitalizations - dependent variables for the logistic regression analyses showing the dichotomies

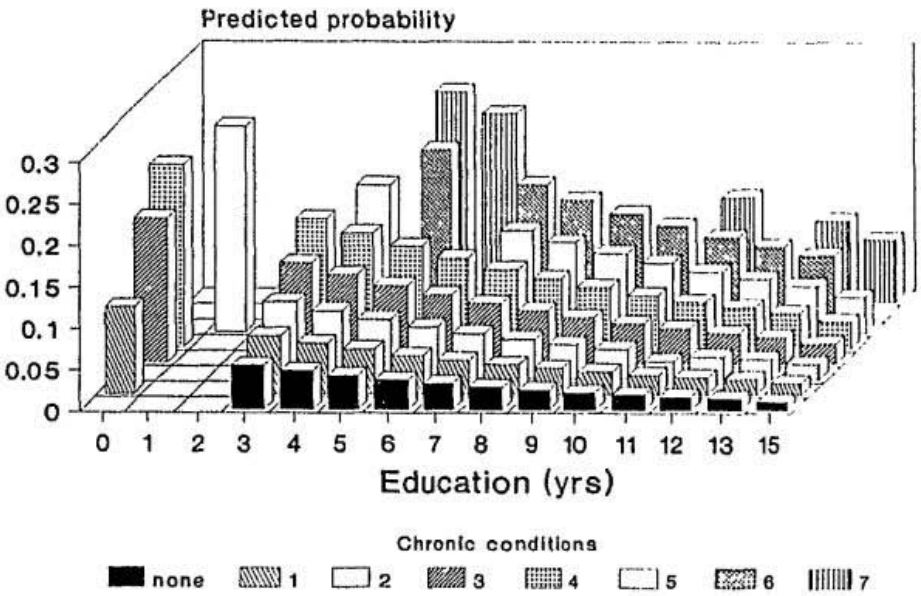
Figures 4.3 to 4.6 show the selected models for REPHOSP. All were good models with the exception of Figure 4.6 where the C.C. Brown chi-square statistic implied that the logistic form was not suitable for the data under consideration. The subjects who were in disagreement on being hospitalized in the recall period were more likely to be less educated and

had a higher number of chronic conditions (Figure 4.3). When those not hospitalized in the recall period were excluded from the agreement group, males were then found to be more likely to disagree than to agree (Figure 4.4). Age was a prominent factor in the occurrence of underreporting; the older population were much more likely to underreport than the younger population (Figure 4.5). When comparing the underreporters to those in agreement the model suggested that age was a relevant variable but the model was not satisfactory (Figure 4.6). No independent variables were selected by the program when comparing the occurrence of overreporting and agreement.

Figures 4.7 to 4.13 show the selected models for the variable ACCHOSP. Under the stricter dichotomy, (i.e. accurate versus less accurate, not very accurate and not accurate) the probability of being inaccurate increased as the subjects reported a higher number of chronic conditions, a lower number of years of education and a lower emotional status (Figure 4.7). Excluding those not hospitalized in the recall period, showed that increased age and being a male, in addition to lower educational status and more chronic conditions showed the most likelihood of being inaccurate when reporting hospitalizations (Figure 4.8). Since SEX was included in this model the models for each sex taken separately were investigated and it was found that increased

education improved the accuracy for males (Figure 4.9), while improvement for females was attributed to a lower number of reported chronic conditions, lower age and higher educational status (Figure 4.10).

Looking at ACCHOSP under a less strict dichotomy showed very similar results with the exception that emotional status and the age of the subject were not important variables in predicting inaccuracy (Figures 4.11 to 4.13).



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: - 2.4572 - 0.1432(EDUC) + 0.2980(CHRCOND)

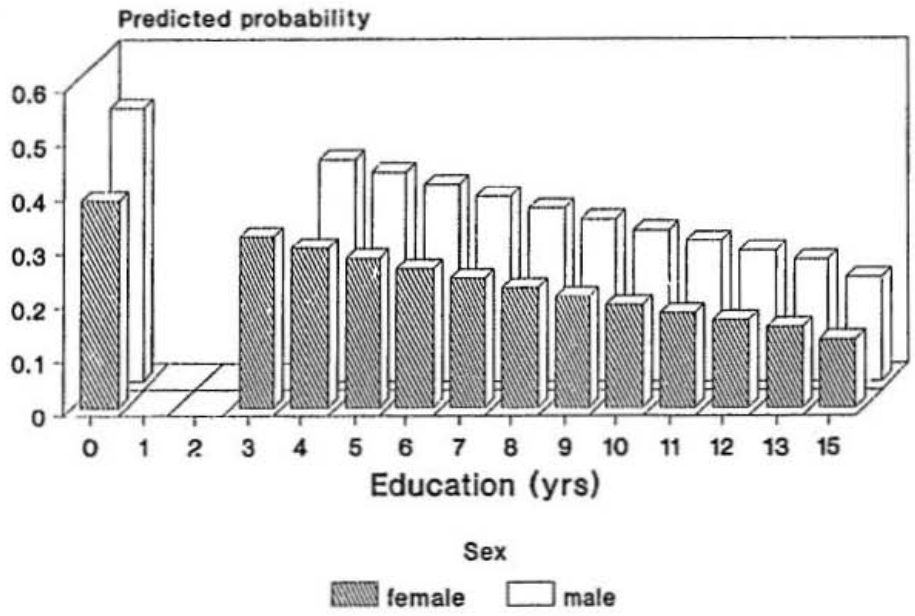
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	3.822	DF = 8	p = 0.873
CC Brown	1.281	DF = 2	p = 0.527

Examples of predicted probabilities:

EDUC (yrs)	CHRCOND (number)	probability
0	0	0.0789
0	7	0.4083
15	0	0.0099
15	7	0.0745

Figure 4.3 Hospitalizations - REPHOSP - predicted probability of disagreement on being hospitalized in recall period



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: 0.0277 - 0.0972(EDUC) - 0.5028(SEX)

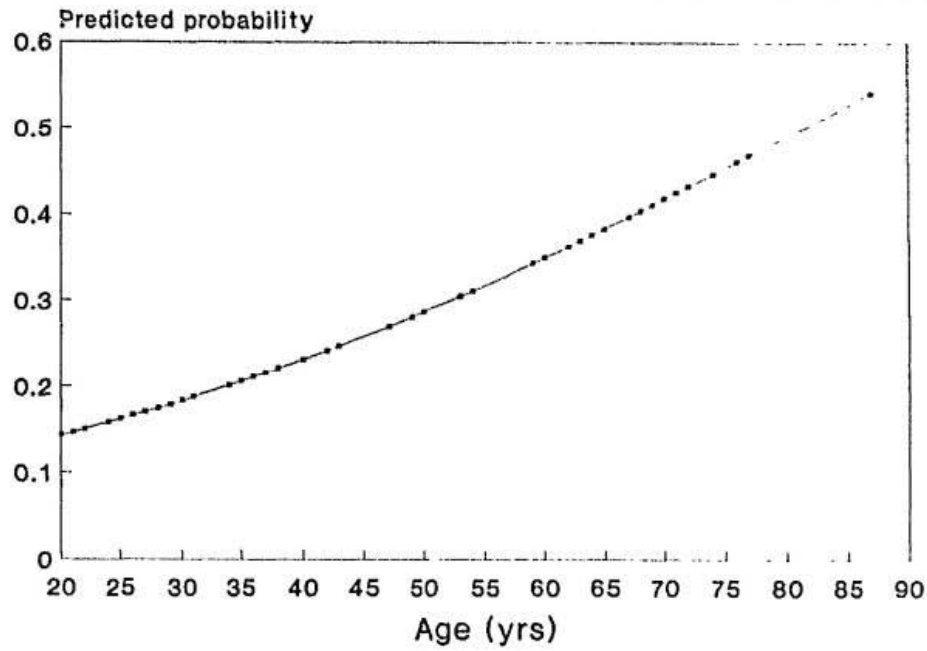
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	5.481	DF = 8	p = 0.705
CC Brown	1.258	DF = 2	p = 0.533

Examples of predicted probabilities:

EDUC (yrs)	SEX	probability
0	F	0.3834
0	M	0.5069
15	F	0.1264
15	M	0.1930

Figure 4.4 Hospitalizations - REPHOSP - predicted probability of disagreement on being hospitalized in recall period, excluding those not hospitalized in the recall period



Model: - 2.5192 + 0.0336(AGE)

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	6.790	DF = 8	p = 0.559
CC Brown	2.619	DF = 2	p = 0.270

Examples of predicted probabilities:

AGE (yrs)	probability
20	0.1361
40	0.2357
60	0.3764
80	0.5415

Figure 4.5 Hospitalizations - REPHOSP - predicted probability of being an under- rather than an overreporter

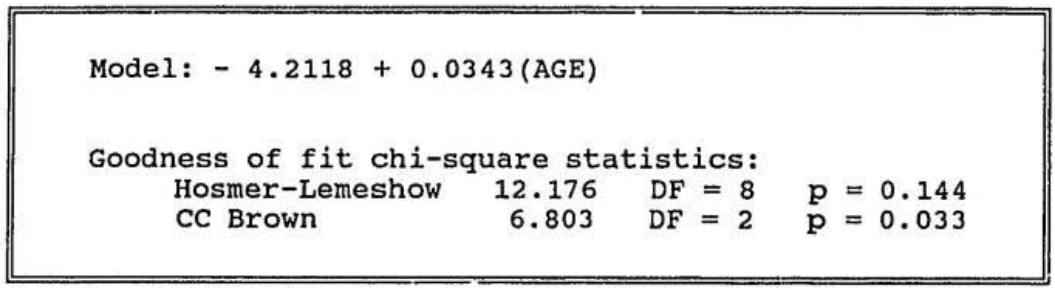
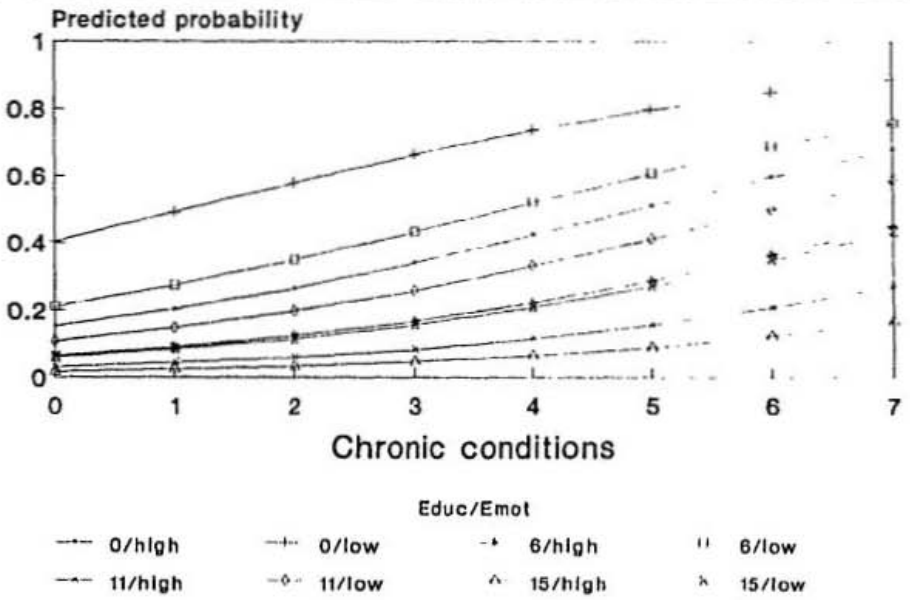


Figure 4.6 Hospitalizations - REPHOSP - predicted probability of being an underreporter rather than in agreement, excluding those not hospitalized in the recall period



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: $-1.7231 - 0.1573(\text{EDUC}) + 0.3530(\text{CHRCOND}) + 0.3083(\text{EMOT}_1) + 0.6429(\text{EMOT}_2) + 1.3361(\text{EMOT}_3)$

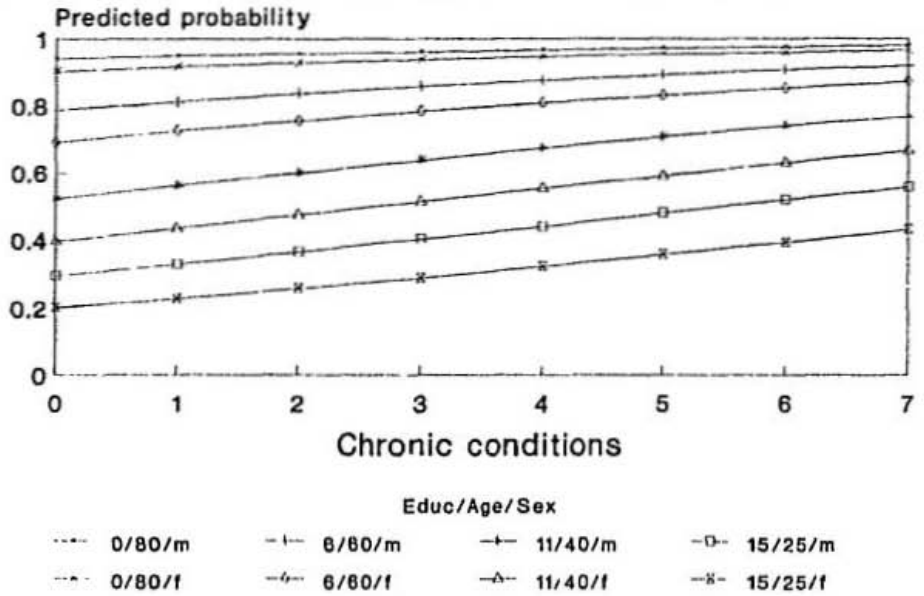
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	2.077	DF = 6	p = 0.912
CC Brown	0.357	DF = 2	p = 0.837

Examples of predicted probabilities:

EDUC (yrs)	CHRCOND (number)	EMOT	probability
0	0	low	0.4044
0	7	low	0.8893
15	0	high	0.0166
15	7	high	0.1664

Figure 4.7 Hospitalizations - ACCHOSP - predicted probability of being inaccurate for hospitalizations in recall period (accurate versus less accurate, not very accurate and not accurate)



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: 1.5133 + 0.0151(AGE) - 0.1841(EDUC) -
0.5076(SEX) + 0.1580(CHRCOND)

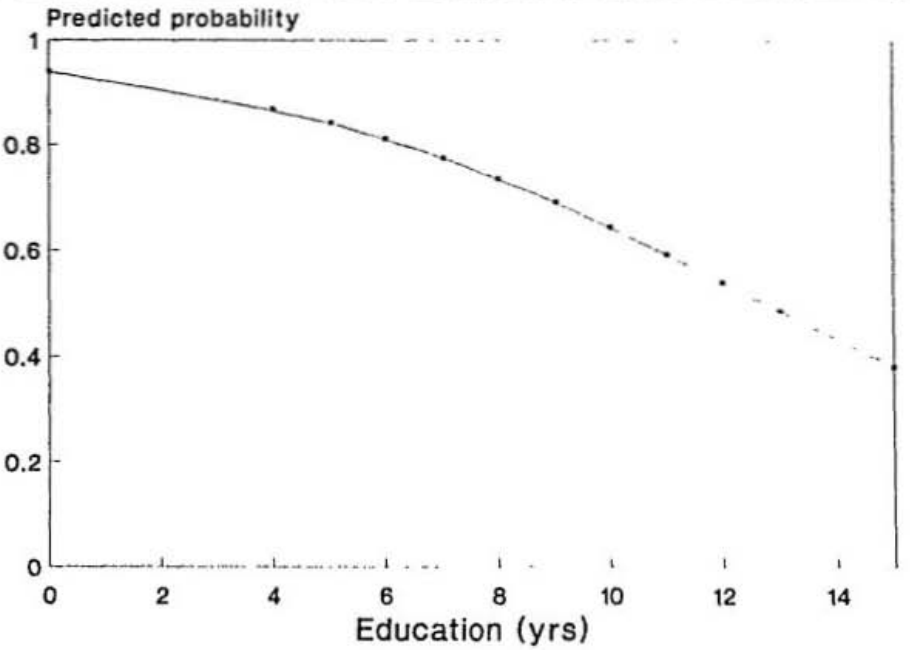
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	6.357	DF = 8	p = 0.607
CC Brown	0.297	DF = 2	p = 0.862

Examples of predicted probabilities:

EDUC (yrs)	AGE (yrs)	SEX	CHRCOND (number)	probability
0	80	M	0	0.9382
0	80	M	7	0.9787
15	25	F	0	0.2012
15	25	F	7	0.4321

Figure 4.8 Hospitalizations - ACCHOSP - predicted probability of being inaccurate for hospitalizations (accurate versus less accurate, not very accurate and not accurate, excluding those not hospitalized in the recall period)



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: 2.7468 - 0.2152(EDUC)

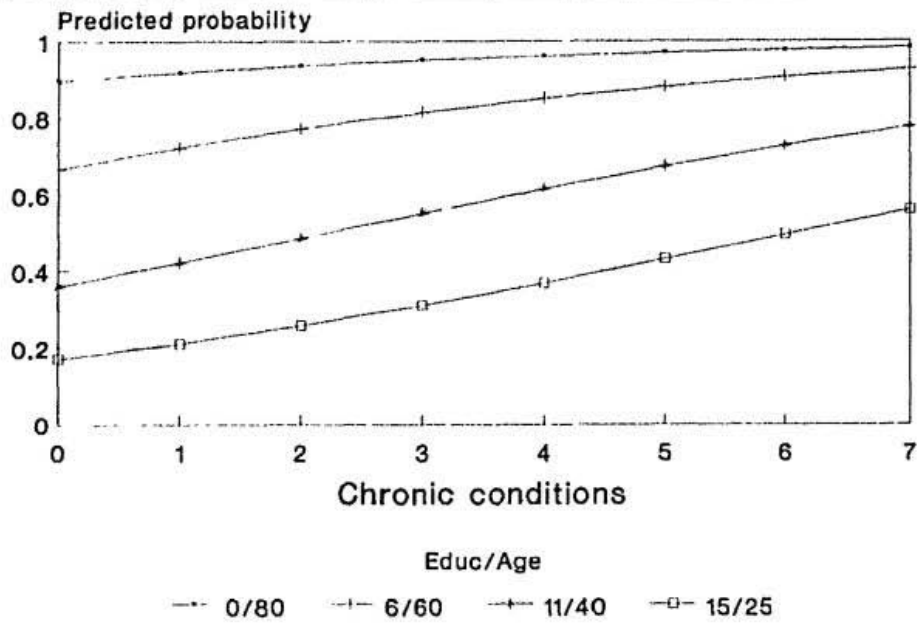
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	6.302	DF = 6	p = 0.390
CC Brown	3.921	DF = 2	p = 0.141

Examples of predicted probabilities:

EDUC (yrs)	probability
0	0.9397
6	0.8109
11	0.5937
15	0.3819

Figure 4.9 Hospitalizations - ACCHOSP - males - predicted probability of being inaccurate for hospitalizations (accurate versus less accurate, not very accurate and not accurate, excluding those not hospitalized in the recall period)



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: 0.7699 + 0.0174(AGE) - 0.1856(EDUC) +
0.2600(CHRCOND)

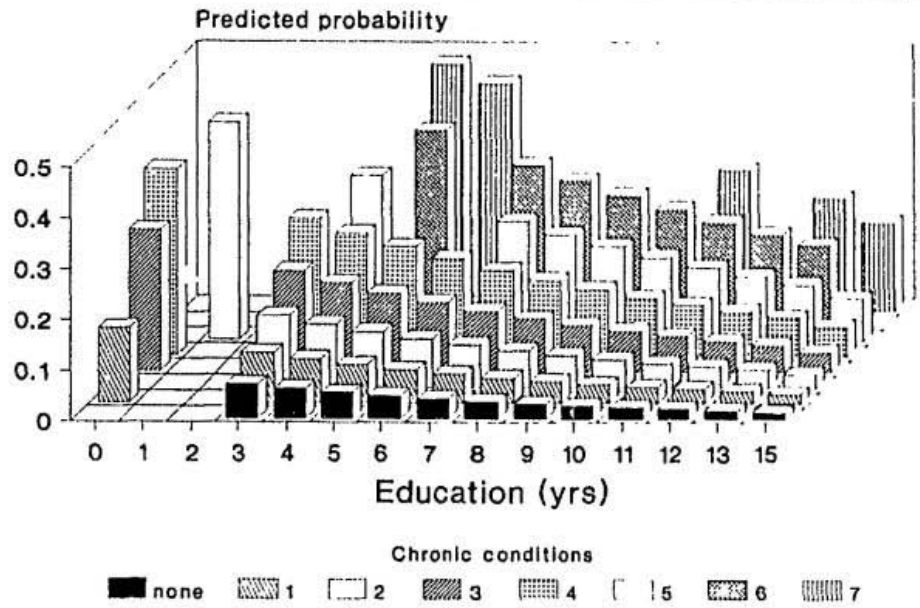
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	4.518	DF = 8	p = 0.808
CC Brown	0.488	DF = 2	p = 0.783

Examples of predicted probabilities:

EDUC (yrs)	AGE (yrs)	CHRCOND (number)	probability
0	80	0	0.8966
0	80	7	0.9817
15	25	0	0.1710
15	25	7	0.5600

Figure 4.10 Hospitalizations - ACCHOSP - females - predicted probability of being inaccurate for hospitalizations (accurate versus less accurate, not very accurate and not accurate, excluding those not hospitalized in the recall period)



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: - 2.1717 - 0.1497(EDUC) + 0.4056(CHRCOND)

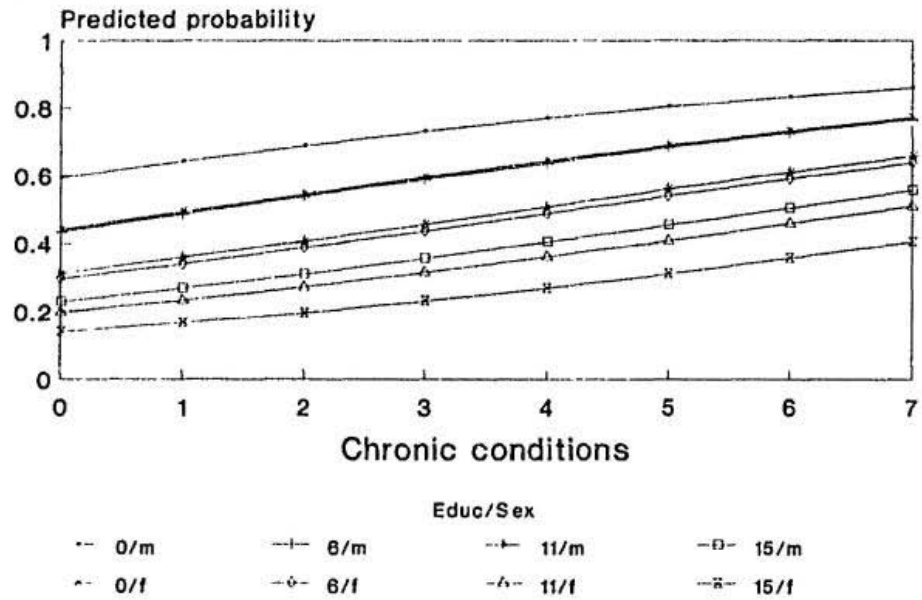
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	5.582	DF = 4	p = 0.233
CC Brown	1.241	DF = 2	p = 0.538

Examples of predicted probabilities:

EDUC (yrs)	CHRCOND (number)	probability
0	0	0.1023
0	7	0.6609
15	0	0.0119
15	7	0.1710

Figure 4.11 Hospitalizations - ACCHOSP - predicted probability of being inaccurate for hospitalizations (accurate and less accurate versus not very accurate and not accurate)



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: 0.3802 - 0.1061(EDUC) - 0.6109(SEX) +
0.2062(CHRCOND)

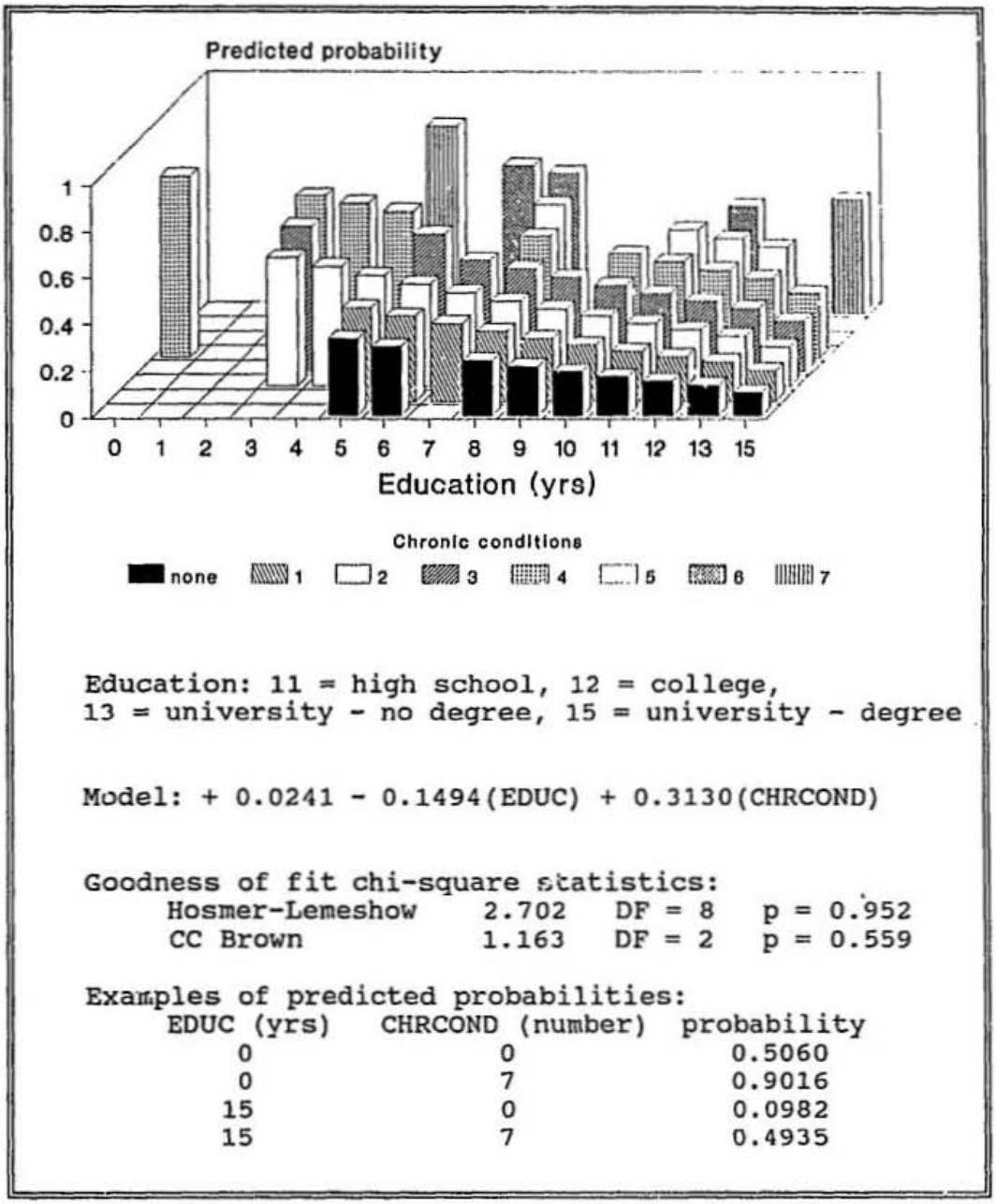
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	3.375	DF = 8	p = 0.909
CC Brown	3.201	DF = 2	p = 0.202

Examples of predicted probabilities:

EDUC (yrs)	SEX	CHRCOND (number)	probability
0	M	0	0.5939
0	M	7	0.8610
15	F	0	0.1392
15	F	7	0.4064

Figure 4.12 Hospitalizations - ACCHOSP - predicted probability of being inaccurate for hospitalizations (accurate and less accurate versus not very accurate and not accurate, excluding those not hospitalized in the recall period)



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: + 0.0241 - 0.1494(EDUC) + 0.3130(CHRCOND)

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	2.702	DF = 8	p = 0.952
CC Brown	1.163	DF = 2	p = 0.559

Examples of predicted probabilities:

EDUC (yrs)	CHRCOND (number)	probability
0	0	0.5060
0	7	0.9016
15	0	0.0982
15	7	0.4935

Figure 4.13 Hospitalizations - ACCHOSP - females - predicted probability of being inaccurate for hospitalizations (accurate and less accurate versus not very accurate and not accurate, excluding those not hospitalized in the recall period) Note: no model was selected for males.

4.2 PHYSICIAN VISITS

4.2.1 RECALL ERRORS: UNDER- AND OVERREPORTING

The average number of visits per year for years 1985 to 1990 inclusive was calculated and compared to each of the years individually. Most subjects had six complete years of physician data, a considerable number only had one or two (these refused longterm access to their data) and a few subjects had died during the course of the six years. Table 4.18 shows the number of these subjects for the various number of years.

Table 4.18
Physician visits - frequency and percentage of subjects with one to six years of data

Number of years of data	Number	Percent
1	331	11.1
2	158	5.3
3	1	0.0
4	1	0.0
5	16	0.5
6	2487	83.1

Table 4.19
Physician visits - number of visits depending on whether the subject was hospitalized in 1985

	Not hospitalized	Hospitalized
NUMBER OF VISITS		
$\bar{x}$	4.1	10.8
SD	5.3	6.8
SE	0.1	0.4
(n)	(2708)	(286)

The range of the differences between the number recorded and the mean value for the six years was from -39 to +28 visits, but for any chosen year over 76% (range 76.5 to 80.7) were within ± 3 visits of the average for 1985-90; the distribution of the differences fitted a normal curve satisfactorily. Therefore, for the large majority of subjects, the number of visits in any year between 1985 and 1990 was close to that of the average of those years. For all further analysis only the year 1985 was considered. Fourteen subjects were unable to remember the number of physician visits made in the recall period so these were excluded from any further analyses.

Table A9 in the Appendix shows the frequency distribution for the number of physician visits according to the official records and Table 4.19 the summary statistics for the subjects divided into those who were hospitalized in the year 1985 and those who were not. There was a significant difference in the number of visits depending on whether the subject was hospitalized or not ($t(324) = -16.05$, $p<0.0001$).

Table 4.20 shows the numbers in the four groups of the variable REPMCP and Table 4.21 shows the overall rates of under- and overreporting together with the Kappa values. Females showed lower rates for both under- and overreporting than males, but the Kappa value was slightly lower indicating

less overall agreement. All Kappa values in Table 4.21 imply moderate agreement.

Table 4.20
Physician visits - comparison of interview report and official MCP records for 1985

Interview report	Official records	Number	%
Agreement			
No visit	No visit	324	10.9
Visit	Visit	2181	73.2
Overreport			
Visit	No visit	234	7.9
Underreport			
No visit	Visit	241	8.1
Total		2980	100.0

Table 4.21
Physician visits - overall rates of under- and overreporting together with the Kappa value, by sex for physician visits

	Male	Female	All
Underreporting (%)	12.9	8.0	10.0
Overreporting (%)	14.1	6.6	9.7
Kappa value	0.49	0.41	0.48

The difference between the reported visits for the year prior to the interview and that obtained from the official records was computed. The range was from -92 to +90 visits, with almost two-thirds of the subjects (74.1%) falling in the -3 to +3 range. Table 4.22 shows the descriptive statistics on this variable. When comparing those subjects who were

hospitalized in the calendar year 1985 against those who were not, there was no significant difference found, even though the means show that those not hospitalized were slightly more inclined to underreport their visits (the mean is negative) than those who were hospitalized (Table 4.23).

Table 4.22
Physician visits - difference between reported and recorded values

$\bar{x}$	-0.492	minimum	-92
SD	5.958	maximum	90
(n)	(2980)	SE	0.109

Table 4.23
Physician visits - difference in reported and recorded values depending on whether the subject was hospitalized in 1985

	Not hospitalized	Hospitalized
DIFFERENCE		
$\bar{x}$	-0.6	0.4
SD	5.0	11.5
SE	0.1	0.7
(n)	(2696)	(284)

For comparison to results from Mechanic and Newton (1965), and Cleary and Jette (1984), the percentages who under- and overreport by three or more visits were calculated for subjects grouped according to the number of visits that were recorded. It was found that 6.2% underreported by three or more visits in the 3-5 visit group, while for 6-7 visits the figure was 29.5% and for eight or more visits, 56.6%

underreport. The corresponding values for overreporting by three or more visits were 5.6% for those who reported 0-2 visits, 9.8% for the 3-5 visit group, 14.7 for the 6-7 group and 13.8 for those with eight or more visits. The percentage of those in agreement declined from 94.4% in the 0-2 visit group to 29.4% in the eight or more group.

Table 4.24
Physician visits - officially recorded number

	Males		Females		Total	
	n	(%)	n	(%)	n	(%)
NUMBER OF VISITS						
0 - 3	898	(66.6)	745	(45.3)	1643	(54.9)
4 - 6	220	(16.3)	367	(22.3)	587	(19.6)
7 - 15	194	(14.4)	421	(25.6)	615	(20.5)
16+	37	(2.7)	112	(6.8)	149	(5.0)

The number of officially recorded visits to physicians could be grouped into four groups where over half had zero to three visits recorded (Table 4.24). As expected, the distributions varied by sex with females being more likely to have made more visits to physicians than males. The mean number of visits for males was 3.6 and 5.8 for females ($t(2991) = -10.73$, $p<0.0001$). Table 4.25 shows that males demonstrated more association with the level of physician utilization than did females. This may have been confounded in females by the presence of visits for pregnancy. Such visits would not be expected to be found at any one education

level more than others, but would be found more in the younger females than the older ones. This, together with the higher use expected in older persons, would result in no clear association being apparent for age.

Table 4.25
Physician visits - association (Gamma) for age and education by level of physician utilization

	Male	Female
AGE	0.32	0.07
EDUCATION	-0.22	-0.13

Table 4.26 shows the trend in age and education for each of the sexes. The range shown in the mean ages was much greater for males than females and this pattern was also apparent for the years of education.

The difference between reported and recorded visits was calculated and Table 4.27 shows the percentage in the various levels of difference and recorded number of visits. This table clearly shows that as the number of recorded visits increases then so did the discrepancy between the number reported and that recorded. These results were looked at in a different perspective, that of under- and overreporting (by any amount) shown in Tables 4.30 later.

Table 4.26
Physician visits - age and education by sex for number of recorded visits (grouped)

	NUMBER OF RECORDED VISITS			
	0 - 3	4 - 6	7 - 15	16+
AGE (yrs)				
MALE				
$\bar{x}$	37.7	41.6	46.1	50.7
SD	13.9	14.3	15.5	16.3
SE	0.5	1.0	1.1	2.7
(n)	(898)	(220)	(194)	(37)
FEMALE				
$\bar{x}$	40.4	42.4	42.1	44.8
SD	15.1	17.1	17.4	18.1
SE	0.5	0.9	0.9	1.7
(n)	(745)	(367)	(421)	(112)
EDUCATION (yrs)				
MALE				
$\bar{x}$	11.7	11.1	10.7	9.6
SD	2.5	2.5	2.8	3.1
SE	0.1	0.2	0.2	0.5
(n)	(896)	(220)	(191)	(37)
FEMALE				
$\bar{x}$	11.4	11.4	10.9	10.4
SD	2.3	2.5	2.5	2.8
SE	0.1	0.1	0.1	0.3
(n)	(743)	(367)	(420)	(111)

Table 4.27
Physician visits - percentage of subjects at various levels of discrepancy for different levels of recorded visits

	NUMBER OF RECORDED VISITS			
	0 - 3	4 - 6	7 - 15	16+
NUMBER OF VISITS DISCREPANCY				
0	36.9	10.1	4.9	2.0
1	35.2	21.0	11.1	2.7
2 - 3	21.6	42.0	21.3	6.8
4 - 6	3.2	21.2	34.5	22.4
7 - 12	2.3	4.4	24.7	27.2
13+	0.7	1.4	3.4	38.8

4.2.2 SOCIO-DEMOGRAPHIC AND HEALTH STATUS VARIABLES

Tables 4.28 and 4.29 show the descriptive statistics for the socio-demographic and health status variables for the four categories of agreement (REPMCP).

The agreement on visits group (visit/visit) showed the most discrepancy of the 4 groups; comparing it to the overreporters, the overreporters were more likely to be male, younger, more educated, had better self-assessed health status, fewer chronic conditions and had better emotional health (SEX: $X^2(1) = 37.82$, $p<0.0001$; AGE: $t(292) = 2.37$, $p<0.05$; EDUCATION: $t(287) = 2.89$, $p<0.01$; SAHS: $t(300) = 4.31$, $p<0.001$; CHRCOND: $t(318) = 7.03$, $p<0.0001$; EMOT: $t(299) = 3.04$, $p<0.01$).

Table 4.28
Physician visits - categories of agreement for MCP information, socio-demographic variables

	Information from interview / Information from records			
	no visit [a] / no visit	visit [a] / visit	visit [b] / no visit	no visit [c] / visit
SEX				
male (%)	69.4	39.0	59.8	52.3
female(%)	30.6	61.0	40.2	47.7
AGE (yrs)				
$\bar{x}$	39.7	41.3	38.8	39.5
SD	14.8	15.9	15.1	14.8
SE	0.8	0.3	1.0	1.0
(n)	(324)	(2181)	(234)	(241)
EDUCATION (yrs)				
$\bar{x}$	11.2	11.3	11.8	11.2
SD	2.5	2.5	2.5	2.1
SE	0.1	0.1	0.2	0.1
(n)	(321)	(2175)	(234)	(241)

[a] Agreement [b] Overreport [c] Underreport

When the underreporters were compared to the agreement on visits group, the underreporters were more likely to be male, had better self-assessed health status, fewer chronic conditions, were more satisfied with their physical condition and had a better emotional score (SEX: $X^2(1) = 15.86$, $p<0.001$; SAHS: $t(313) = 5.33$, $p<0.0001$; CHRCOND: $t(348) = 9.76$, $p<0.0001$; PHYSCOND: $t(306) = 3.34$, $p<0.001$; EMOT: $t(309) = 2.62$, $p<0.01$). AGE was not significant ($t(305) = 1.82$, $p<0.10$).

Table 4.29
Physician visits - categories of agreement for MCP information, health status variables

	Information from interview / Information from records			
	no visit [a] / no visit	visit [a] / visit	visit [b] / no visit	no visit [c] / visit
SAHS				
$\bar{x}$	1.8	2.0	1.8	1.7
SD	0.7	0.7	0.6	0.6
SE	0.04	0.02	0.04	0.04
(n)	(324)	(2181)	(234)	(241)
CHRCOND				
$\bar{x}$	0.4	1.2	0.7	0.6
SD	0.7	1.3	1.0	0.9
SE	0.04	0.03	0.07	0.06
(n)	(324)	(2181)	(234)	(241)
PHYSCOND				
$\bar{x}$	1.9	2.0	2.0	1.9
SD	0.6	0.6	0.7	0.6
SE	0.04	0.01	0.04	0.04
(n)	(324)	(2180)	(234)	(241)
EMOT				
$\bar{x}$	1.5	1.7	1.6	1.6
SD	0.6	0.7	0.6	0.6
SE	0.03	0.02	0.04	0.04
(n)	(324)	(2173)	(234)	(240)

[a] Agreement [b] Overreport [c] Underreport

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

4.2.3 EXPLANATORY ANALYSIS

The accuracy of the number of visits reported is shown in Tables 4.30 and 4.31. In Table 4.30 a discrepancy of ±1 visit was allowed in the category called agreement. This table shows that the more visits made to physicians the more likely the subject was to underreport the number. Overreporting did not decrease in proportion to the increase seen in underreporting, instead it was the agreement group which decreased as the underreporters increased. Table A10 shows the percentages under a stricter categorization where agreement implied the same number of visits was reported and recorded. Table 4.31 shows the frequency by sex for the summary variable ACCMCP.

Table 4.30
Physician visits - percentage of subjects at various levels of accuracy for different levels of recorded visits

	NUMBER OF RECORDED VISITS				
	0 - 3	4 - 6	7 - 15	16+	Total
ACCURACY OF REPORTING					
Underreporting ≥2 visit	8.4	51.7	64.0	76.2	31.7
Agreement ± 1 visit	72.1	31.1	16.0	4.8	49.2
Overreporting ≥2 visit	19.4	17.2	20.0	19.0	19.1

Table 4.31 Physician visits - summary accuracy variable (ACCMCP), by sex				
	Male		Female	
	n	%	n	%
Accurate	754	56.2	713	43.5
Less accurate	412	30.7	522	31.9
Not accurate	176	13.1	403	24.6
(n)	(1342)	(100)	(1638)	(100)

Table 4.32 Physician visits - associations (Gamma) between socio-demographic, health status variables and accuracy (ACCMCP)			
	All (n=2980)	Male (n=1342)	Female (n=1638)
SEX	0.26	-	-
AGE (grouped)	0.14	0.26	0.05
EDUC (grouped)	-0.13	-0.14	-0.11
SAHS	0.21	0.19	0.23
CHRCOND	0.30	0.36	0.24
PHYSCOND	0.07	0.05	0.10
EMOT	0.13	0.11	0.13

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

Tables 4.32 and 4.33 show the association and correlation of the physician visit accuracy variable (ACCMCP) with the socio-demographic and health status variables.

There was reasonable association with SEX, SAHS and CHRCOND, but minimal association with any of the other variables. The Spearman correlation values for ACCMCP (Table 4.33) correspond to the associations found in Table 4.32. As no new associations were unveiled for ACCMCP as compared to those found for ACCHOSP, the same independent variables were used in the logistic regression analyses for physician visits as were utilized for hospitalizations. The variables with their dichotomies used are shown in Figure 4.14.

REPMCP	- agreement versus under- and overreporters
	- underreporters versus overreporters
	- underreporters versus those in agreement
	- overreporters versus those in agreement
ACCMCP	- accurate versus less accurate and not accurate
	- accurate and less accurate versus the not accurate

Figure 4.14 Physician visits - dependent variables for the logistic regression analyses showing the dichotomies to be used

Table 4.33
Physician visits - correlations between variables under consideration (coefficient and significance level)

	AGE (yrs)	EDUC (yrs)	SEX	SAHS	CHRCOND	PHYSCOND	EMOT
EDUC (yrs)	-.3409 .000						
SEX	.0487 .004	-.0455 .006					
SAHS	.1532 .000	-.2489 .000	.0011 .475				
CHRCOND	.3082 .000	-.1500 .000	.1384 .000	.3260 .000			
PHYSCOND	-.0762 .000	.0649 .000	-.0169 .178	.3548 .000	.1763 .000		
EMOT	.0810 .000	-.0498 .003	.0383 .018	.2183 .000	.1503 .000	.2102 .000	
ACCMCP	.0859 .000	-.1063 .000	.1505 .000	.1440 .000	.2314 .000	.0429 .010	.0846 .000

NB the first 6 rows are the same as those in Table 4.16, just ACCMCP row is new. Differences in the 6 rows are due to slightly different numbers of subjects.

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status
ACCMCP: accuracy of reporting physician visits

The logistic regression models for REPMCP are shown in Figures 4.15 to 4.23. For physician visits, the main discriminating variables for being in disagreement were SEX, SAHS and CHRCOND. Those subjects more likely to be in disagreement were males, those with excellent health status and those with a low number of chronic conditions (Figure 4.15). Performing the regression analysis separately by SEX did not produce an acceptable model for either sex, but CHRCOND was implicated as the major contributing variable for both sexes (Figure 4.16). The variables CHRCOND and EDUCATION discriminated between the under- and the overreporters; the underreporters were more likely to be those individuals who had the least years of education and the least number of chronic conditions (Figure 4.17).

Underreporters could not be reliably distinguished from those in agreement (Figure 4.18) since the C.C. Brown statistic is small, but the model indicates that education level, sex, self-assessed health status and the number of chronic conditions were all possible confounders for underreporting. Looking separately by SEX the models were acceptable, for males CHRCOND was the only selected variable showing that the higher the number of reported conditions the more likely they were to be in agreement on having made visits (Figure 4.19); for females, EDUCATION, CHRCOND and SAHS were all important variables, with those in agreement

more likely to have had poor SAHS, a higher grade of education and many chronic conditions (Figure 4.20).

Comparing the overreporters to those in agreement, the overreporters were more likely to be those with a high number of years of education, a low number of chronic conditions and they were more likely to be male than female (Figure 4.21). When looked at by SEX, there was no suitable model produced for males (Figure 4.22) while in females, CHRCOND was an important variable where the more conditions an individual reported the more likely they were to be in agreement (Figure 4.23).

The accuracy of reporting physician visits (ACCMCP) is shown in Figures 4.24 to 4.28. When comparing the accurate to the combined group of less accurate and not accurate, no suitable model was selected for either all subjects together or for males on their own (Figure 4.24). Analysis of females cases resulted in a suitable logistic model, where those with poor health status, low education and a high number of chronic conditions were more inclined to be inaccurate for reporting visits (Figure 4.25).

When the dichotomy of being accurate was relaxed to include the less accurate, leaving the not accurate in a group on their own; poor SAHS, poor emotional status, high number of chronic conditions and being female were all indicators of a subject being more prone to inaccuracy

(Figure 4.26). Looking at males on their own, the model selected was not adequate for the data (Figure 4.27); and for females, poor SAHS with a high number of chronic conditions was the worse combination for being inaccurate in reporting physician visits (Figure 4.28).

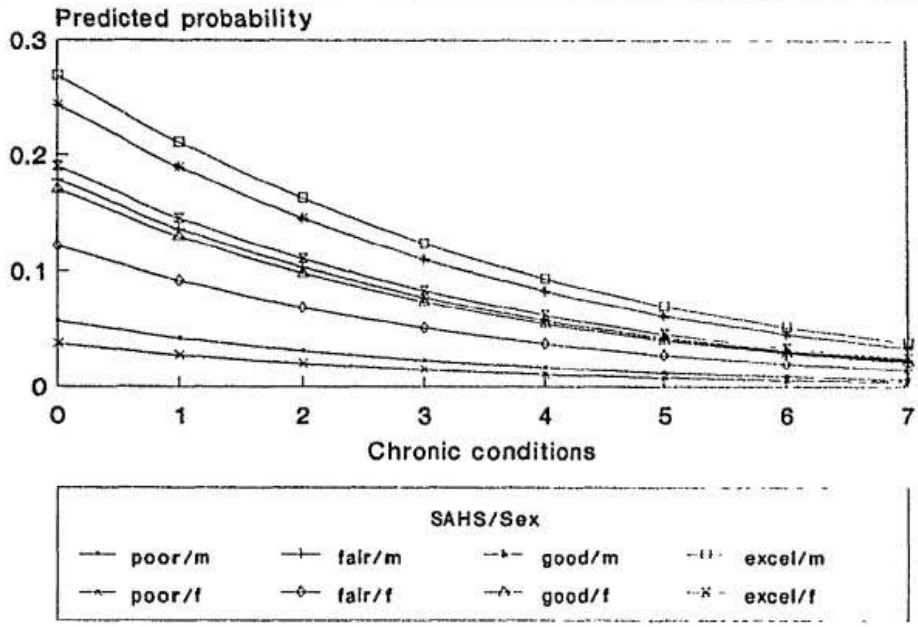
4.3 COMPARISON OF ACCURACY OF HOSPITALIZATIONS AND PHYSICIANS

Table 4.34 shows the association levels when comparing the variables ACCHOSP with ACCMCP. The associations were increased when those subjects not hospitalized in the recall period were excluded. Males showed a stronger association than females.

Table 4.34
Associations (Gamma) between the accuracy variables for hospitalizations and physician visits

	Male	Female	All
ACCHOSP and ACCMCP	0.58	0.32	0.44
ACCHOSP and ACCMCP - excluding those not hospitalized in the recall period	0.63	0.36	0.50

There was no significant association between REPHOSP and REPMCP using chi-square as the measure of association, either when including or excluding those subjects not hospitalized in the recall period.



Model: $-1.0026 - 0.4489(SEX) - 0.3183(CHRCOND) - 0.1331(SAHS_1) - 0.5261(SAHS_2) - 1.8095(SAHS_3)$

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	5.563	DF = 8	p = 0.696
CC Brown	1.125	DF = 2	p = 0.570

Examples of predicted probabilities:

SAHS	SEX	CHRCOND (number)	probability
poor	F	0	0.9631
poor	F	7	0.9956
excellent	M	0	0.7316
excellent	M	7	0.9593

Figure 4.15 Physicians - REPMCP - predicted probability of disagreement on visiting physicians in recall period

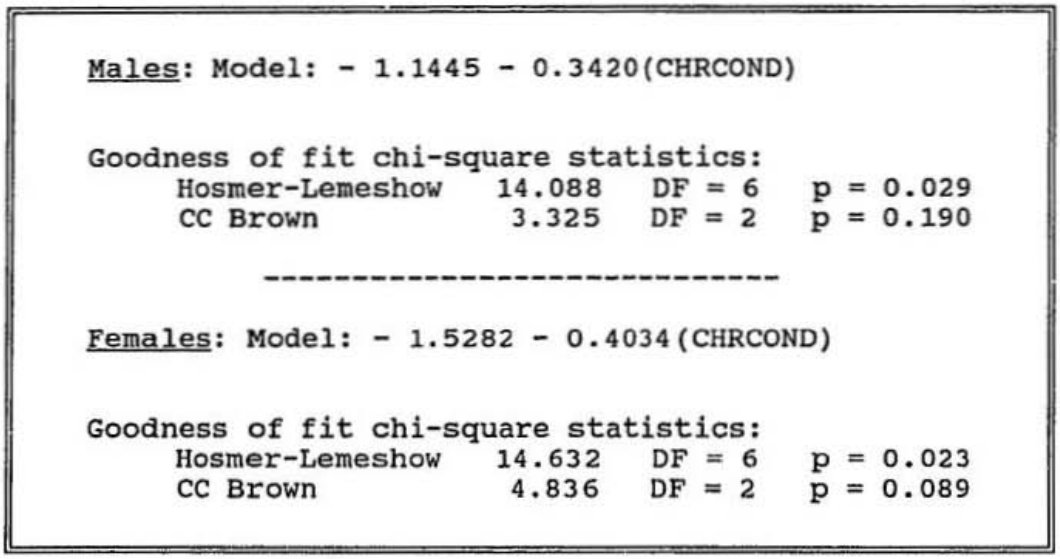
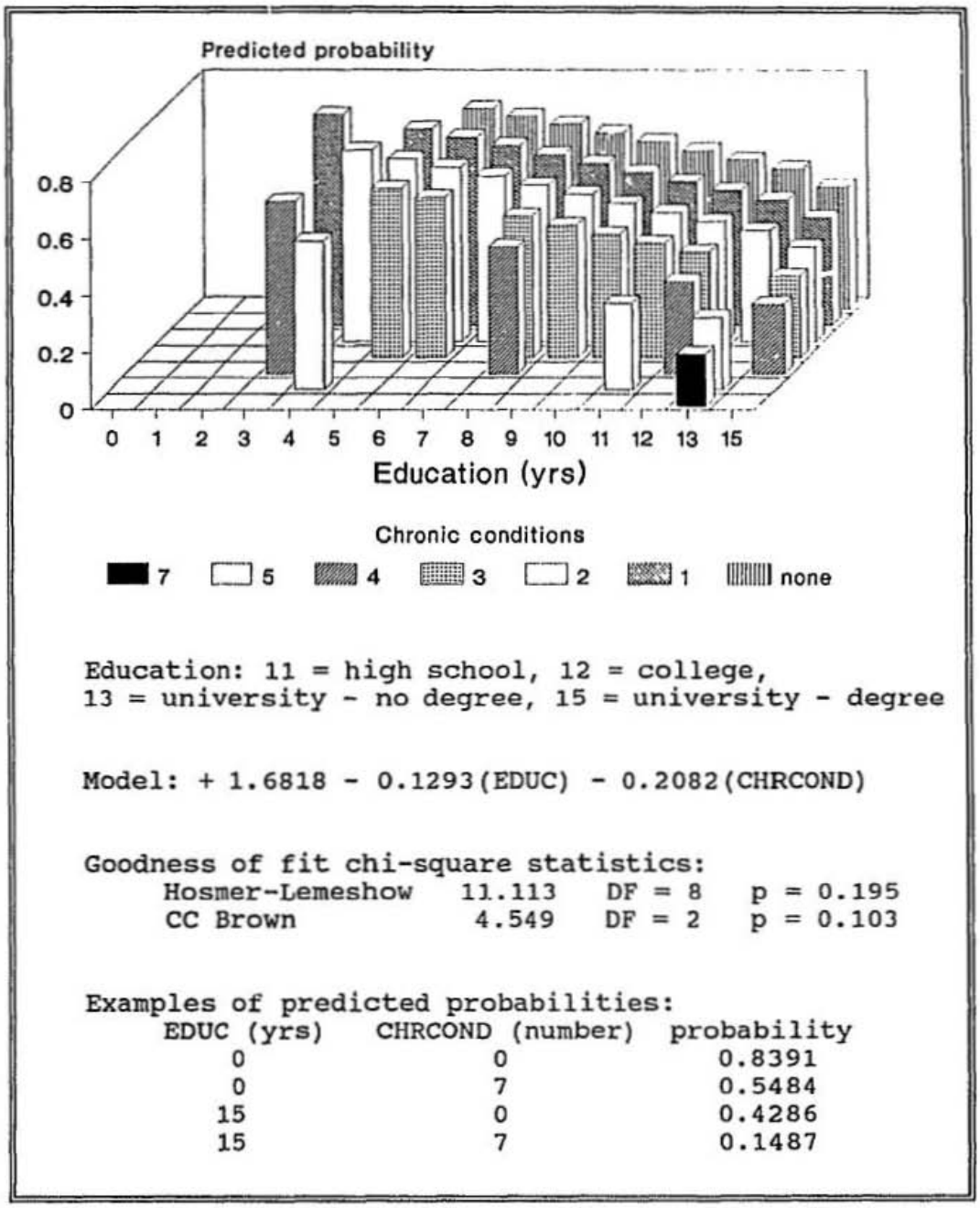


Figure 4.16 Physicians - REPMCP - by sex - logistic regression models for predicted probability of disagreement on visiting physicians in recall period



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: + 1.6818 - 0.1293(EDUC) - 0.2082(CHRCOND)

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	11.113	DF = 8	p = 0.195
CC Brown	4.549	DF = 2	p = 0.103

Examples of predicted probabilities:

EDUC (yrs)	CHRCOND (number)	probability
0	0	0.8391
0	7	0.5484
15	0	0.4286
15	7	0.1487

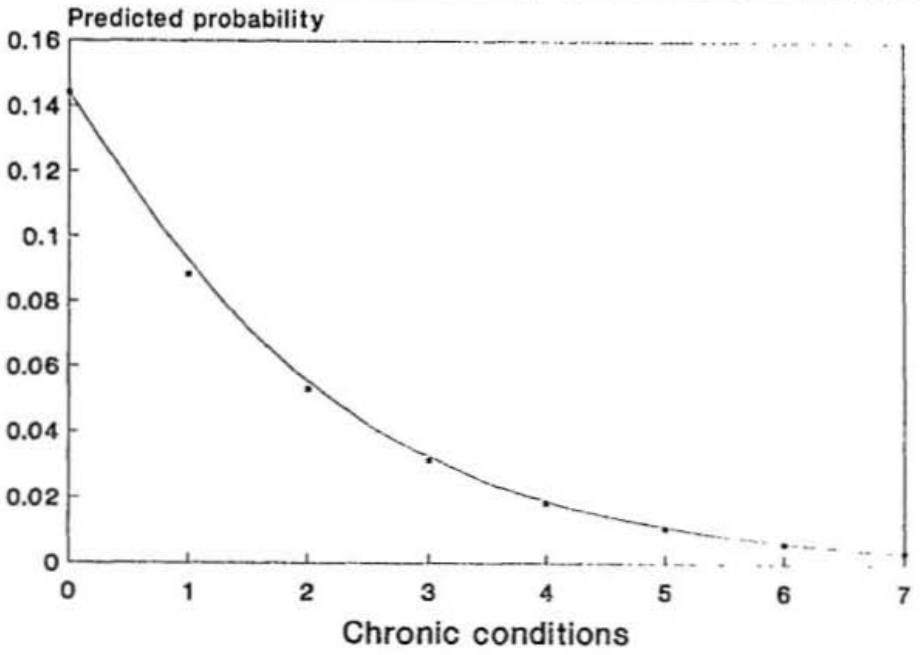
Figure 4.17 Physicians - REPMCP - predicted probability of being an under- rather than an overreporter

Model: $-0.8641 - 0.0686(EDUC) - 0.2871(SEX) - 0.1746(SAHS_1) - 0.6123(SAHS_2) - 14.435(SAHS_3) - 0.4361(CHRCOND)$

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	10.637	DF = 8	p = 0.223
CC Brown	7.310	DF = 2	p = 0.026

Figure 4.18 Physicians - REPMCP - logistic regression model for predicted probability of being an underreporter rather than in agreement



Model: - 1.7825 - 0.5523(CHRCOND)

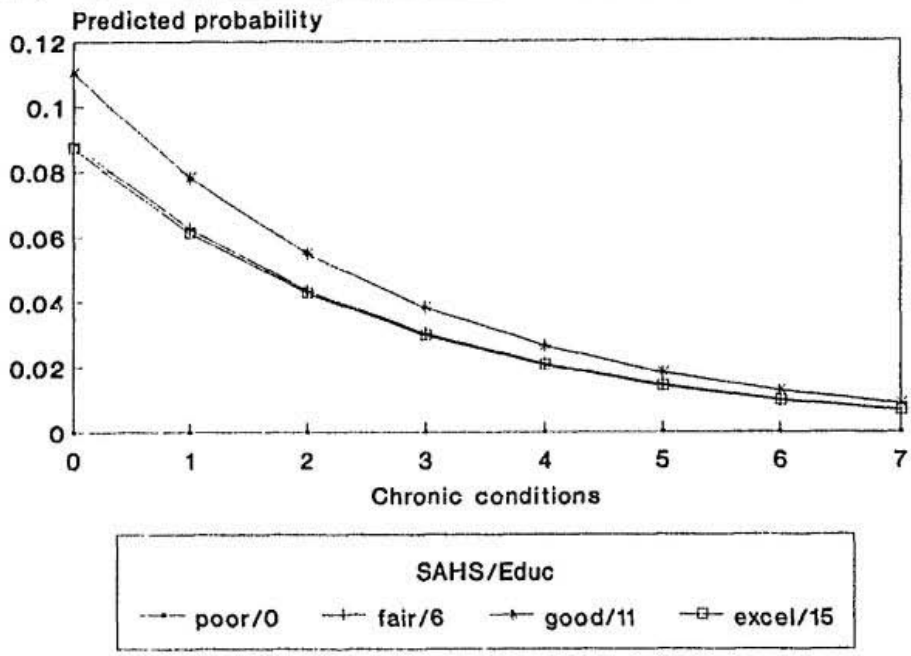
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	2.192	DF = 4	p = 0.700
CC Brown	2.492	DF = 2	p = 0.288

Examples of predicted probabilities:

CHRCOND	probability
0	0.9397
7	0.8109

Figure 4.19 Physicians - REPMCP - males - predicted probability of being an underreporter rather than in agreement



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: $-0.6848 - 0.1109(EDUC) - 0.3810(CHRCOND) - 0.1829(SAHS_1) - 0.9791(SAHS_2) - 10.471(SAHS_3)$

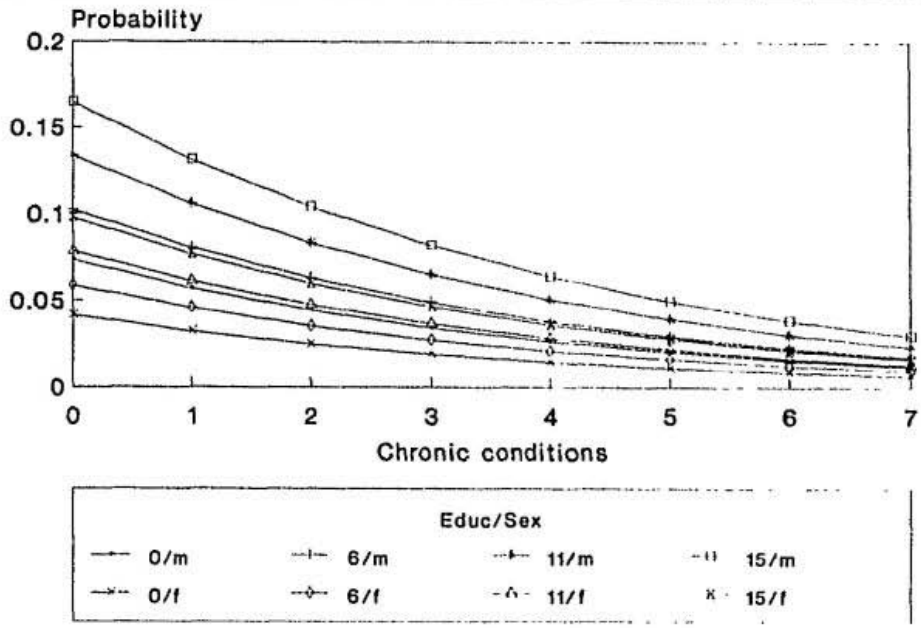
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	12.686	DF = 8	p = 0.123
CC Brown	1.805	DF = 2	p = 0.406

Examples of predicted probabilities:

SAHS	EDUC	CHRCOND (number)	probability
good	11	0	0.8897
good	11	7	0.9915
poor	0	0	1.0000
poor	0	7	1.0000

Figure 4.20 Physicians - REPMCP - females - predicted probability of being an underreporter rather than in agreement



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: - 2.5397 + 0.0611(EDUC) - 0.6010(SEX) -
0.2658(CHRCOND)

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	7.663	DF = 8	p = 0.467
CC Brown	0.069	DF = 2	p = 0.966

Examples of predicted probabilities:

EDUC (yrs)	SEX	CHRCOND (number)	probability
0	F	0	0.9585
0	F	7	0.9933
15	M	0	0.8353
15	M	7	0.9702

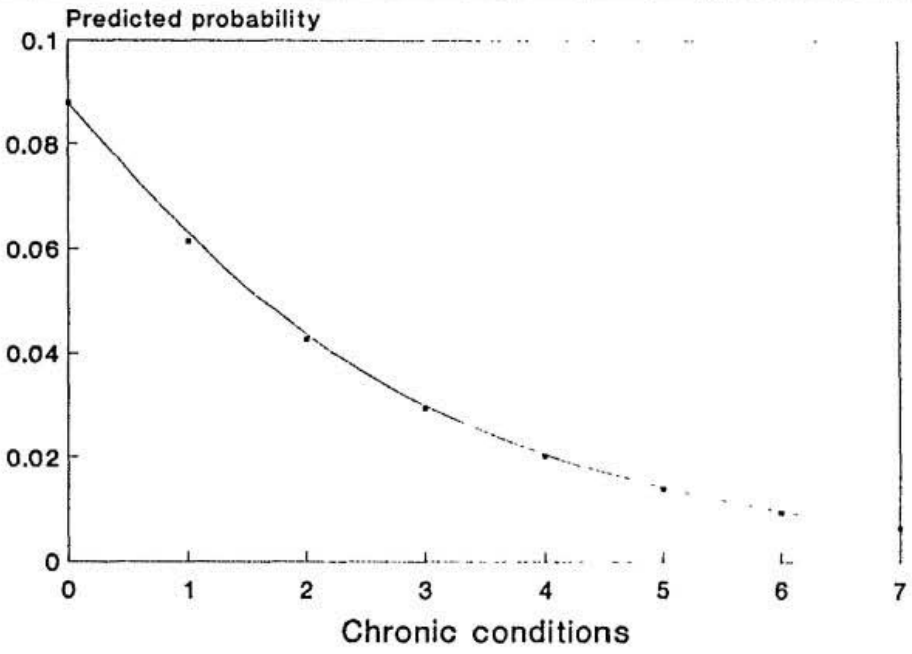
Figure 4.21 Physicians - REPMCP - predicted probability of being an overreporter rather than in agreement

Model: $-1.3131 - 0.0120(AGE) - 0.2414(SAHS_1) - 0.7121(SAHS_2) - 15.357(SAHS_3)$

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	13.795	DF = 8	p = 0.087
CC Brown	0.000	DF = 0	p = 1.000

Figure 4.22 Physicians - REPMCP - males - logistic regression model for predicted probability of being an overreporter rather than in agreement



Model: - 2.3391 - 0.3867(CHRCOND)

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	2.228	DF = 4	p = 0.694
CC Brown	0.000	DF = 0	p = 1.000

Examples of predicted probabilities:

CHRCOND	probability
0	0.0879
7	0.0064

Figure 4.23 Physicians - REPMCP - females - predicted probability of being an overreporter rather than in agreement

All subjects:
Model: $-0.1175 - 0.0432(EDUC) + 0.4323(SEX) + 0.0655(SAHS_1) + 0.2307(SAHS_2) + 1.1758(SAHS_3) + 0.2661(CHRCOND) - 0.0025(EMOT_1) + 0.3182(EMOT_2) + 1.1092(EMOT_3)$

Goodness of fit chi-square statistics:

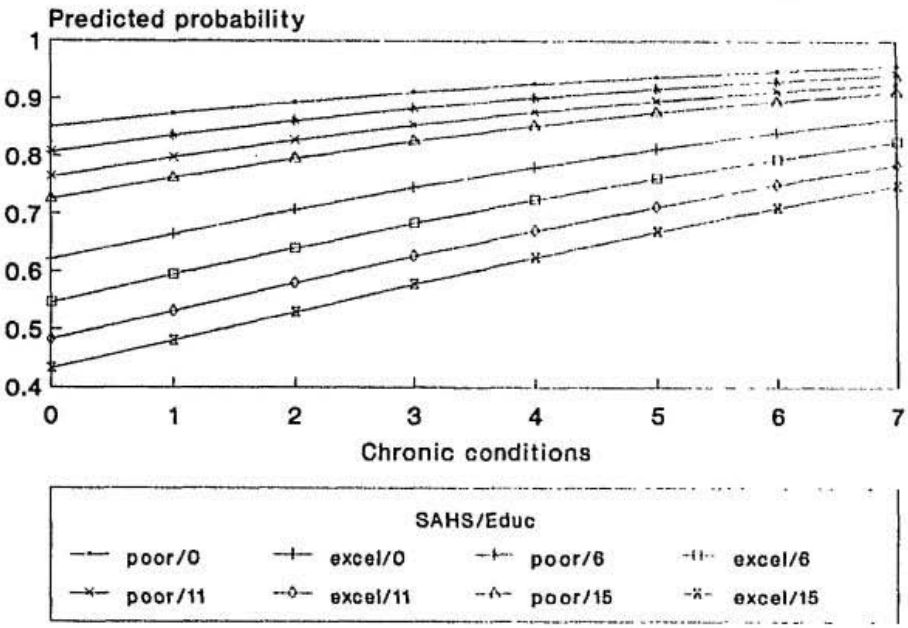
Hosmer-Lemeshow	15.584	DF = 8	p = 0.049
CC Brown	5.321	DF = 2	p = 0.070

Males:
Model: $-1.0473 + 0.0108(AGE) + 0.3816(CHRCOND) - 0.0513(EMOT_1) + 0.5766(EMOT_2) + 1.8581(EMOT_3)$

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	8.750	DF = 8	p = 0.364
CC Brown	4.979	DF = 2	p = 0.083

Figure 4.24 Physicians - ACCMCP - all subjects and males - logistic regression models for predicted probability of being inaccurate for visits (accurate versus less accurate and not accurate)



Education: 11 = high school, 12 = college,
13 = university - no degree, 15 = university - degree

Model: + 0.4914 - 0.0508(EDUC) + 0.1947(CHRCOND) +
$0.0684(SAHS_1) + 0.3892(SAHS_2) + 1.2417(SAHS_3)$

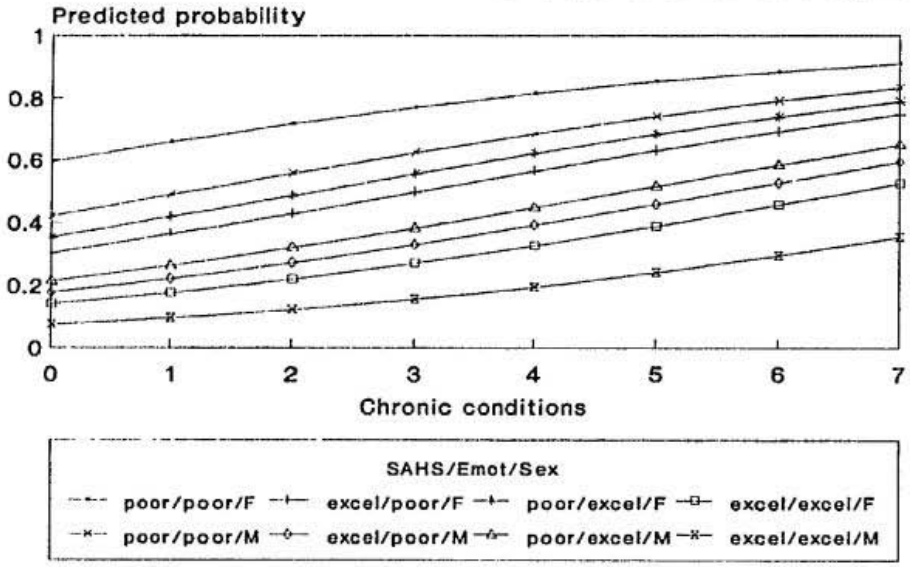
Goodness of fit chi-square statistics:

Hosmer-Lemeshow	9.561	DF = 8	p = 0.297
CC Brown	0.882	DF = 2	p = 0.643

Examples of predicted probabilities:

SAHS	EDUC	CHRCOND (number)	probability
poor	0	0	0.8498
poor	0	7	0.9567
excellent	15	0	0.4328
excellent	15	7	0.7488

Figure 4.25 Physicians - ACCMCP - females - predicted probability of being inaccurate for visits (accurate versus less accurate and not accurate)



Model: $- 2.5298 + 0.7036(SEX) + 0.2742(CHRCOND) + 0.3108(SAHS_1) + 0.6792(SAHS_2) + 1.2203(SAHS_3) - 0.0663(EMOT_1) + 0.2651(EMOT_2) + 0.9864(EMOT_3)$

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	8.126	DF = 8	p = 0.421
CC Brown	3.906	DF = 2	p = 0.142

Examples of predicted probabilities:

SAHS	EMOT	SEX	CHRCOND (number)	probability
poor	poor	F	0	0.5940
poor	poor	F	7	0.9089
excel	excel	M	0	0.0738
excel	excel	M	7	0.3520

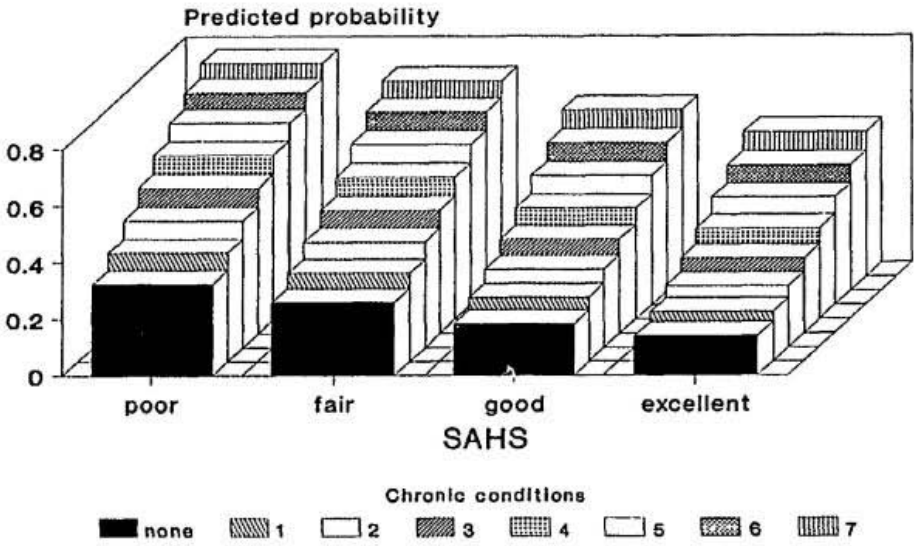
Figure 4.26 Physicians - ACCMCP - predicted probability of being inaccurate for visits (accurate and less accurate versus not accurate)

Model: $-3.2209 + 0.0187(AGE) + 0.2320(CHRCOND) + 0.2595(SAHS_1) + 0.5225(SAHS_2) + 1.2561(SAHS_3) - 0.1274(EMOT_1) + 0.3741(EMOT_2) + 1.8872(EMOT_3)$

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	11.329	DF = 8	p = 0.184
CC Brown	7.892	DF = 2	p = 0.019

Figure 4.27 Physicians - ACCMCP - males - logistic regression model for predicted probability of being inaccurate for visits (accurate and less accurate versus not accurate)



Model: $- 1.8176 + 0.2648(CHRCOND) + 0.3211(SAHS_1) + 0.7651(SAHS_2) + 1.0843(SAHS_3)$

Goodness of fit chi-square statistics:

Hosmer-Lemeshow	2.639	DF = 8	p = 0.916
CC Brown	3.730	DF = 2	p = 0.155

Examples of predicted probabilities:

SAHS	CHRCOND (number)	probability
poor	0	0.3245
poor	7	0.7540
excellent	0	0.1397
excellent	7	0.5090

Figure 4.28 Physicians - ACCMCP - females - predicted probability of being inaccurate for visits (accurate and less accurate versus not accurate)

5 DISCUSSION

The most significant finding from this analysis is the reassuring fact that people in general are mostly accurate when reporting hospitalizations that have occurred in a 12 month period immediately prior to an interview. For the whole population surveyed, 97.3% gave the correct information on such hospitalizations. When the non-hospitalized are excluded from the survey group, the percentage falls to 78.6% which equates to only one in four hospitalized subjects giving erroneous information.

The values for under- and overreporting of 7.3% and 16.2% for hospitalizations are similar to those found by Schach *et al.* (1972) and Belloc (1954); there are slightly fewer underreports and more overreports than either of these groups found. The study by Schach *et al.* is the most comparable to this study since it was carried out in Saskatchewan where a universal medical insurance plan is in effect and all hospitalizations are recorded by the Provincial government. Their values of 11% underreports and 14% overreports differ slightly from those found here. The percentage in agreement for those where hospitalizations were reported is the same in the two studies, 78% for Schach *et al.* and 78.6% in St. John's.

The differences in under- and overreporting hospitalizations by sex is worth noting; females are much

less likely to both under- and overreport than males. This is contrary to Cannell *et al.* (1977) who found minimal differences between the sexes for either hospital or physician visits. The difference found in this survey may be in part due to the fact that females are much more likely to go to hospital than males (because of pregnancy and related conditions) and hence they have more events to recall.

For physician visits, the observed agreement is not so marked, possibly because 1) the occurrence of a visit to a physician is not so memorable as a hospitalization, and 2) in our case, the data are less robust than they are for hospitalizations. 84.1% of the total surveyed population gave correct information on physician visits. Underreports of 10.0% and overreports of 9.7% for physician visits is much less than those found by Andersen *et al.* (1979) and Jobe *et al.* (1990). Both these two comparison papers deal with the US population where there is no universal medical insurance and where the verification of data is much more complex and often impractical to complete, even though the recall of visits by an individual may be more thorough since payment should help in making the event more memorable.

As with hospitalizations, females have a lower value for both under- and overreporting physician visits than males. The reasons for this are not so apparent as for hospitalizations since nearly all subjects visit the

physician at least once a year. One possible reason is that males may be more likely to suppress the information and hence underestimate the number of visits since they do not place as much significance on them as do females. The overall Kappa value of 0.48 for physician visits implies only moderate agreement.

Analyses like those undertaken by Yaffe *et al.* (1978), Green *et al.* (1979) and Coulter *et al.* (1985) on the type of physician/hospital visit or the occurrence of a surgical procedure during hospitalization are not possible on the data collected for this analysis. The data for hospitalizations do not contain any information on surgical procedures and the physician visit data do not contain any code for place or reason of visit. Diagnoses for hospitalizations are only available for those recorded in the official database (those in agreement and the underreporters) and the number at disagreement, where diagnoses are available, (i.e. the underreporters - 23) is too small to make any useful analysis by diagnosis.

It has been shown that physician underreporting is related to the number of visits that occurred (Mechanic and Newton, 1965; Cleary and Jette, 1984). This is apparent in this analysis too, in that for the 3-5 visit group 6.2% underreport by three or more visits, while for patients with eight or more visits, 56.6% underreport the number. These

values are considerably less than those found by Mechanic and Newton (1965) but their data are restricted to male college students so are not directly comparable. The corresponding values for overreporting by three or more visits range from 5.6% for those who reported 0-2 visits to a high of 14.7% for those in the 6-7 visit group.

Andersen *et al.* (1979) had found that infrequent users of physician services tended to underreport while those who used the services more tended to overreport; this is not apparent for the St. John's data where the percentage underreporting by two or more visits rises dramatically with the number of visits recorded while there is no trend to be seen for the overreporters (Table 4.30). When the tolerance in the number of visits is reduced to one, the percentage underreporting ranges from 24.9% to 77.6%, and the overreporting from 38.1% down to 20.4%.

Comparing the percentages who underreport for various numbers of hospital stays these data show that 7.7% of those with one stay underreport, and 5.9% of those with three or more visits underreport; this is less than the values of 17% and 24% found in one study in the USA (NCHS, 1965). The lower values seen in these data could be due to the fact that the subjects have to think of fewer providers and only one insurance scheme for hospitalizations so lapses in memory may be less than subjects living in large cities in the USA. The

study by the NCHS sampled known hospital users from a group of hospitals and excluded any female who was admitted only for normal delivery in the reference year. They state that pregnancy visits are mostly recalled correctly so by excluding them one would expect to find higher rates of underreporting. Many studies have found that the longer the length of stay the more likely the hospitalization is to be reported and that more underreporting occurs with shorter length of stay; this is not apparent in our data. There is a trend for higher underreporting in the shorter lengths of stay (1-2 days), but it as high in those hospitalized for over 15 days. The percentages of agreements also follow a trend for the groups under 15 days, but the figure for 15 and over does not follow the trend.

The same length of stay as that given by the hospitalization records was reported by 27.6% of subjects, 29.6% recalled a shorter stay than the verification data and 42.8% recalled more. The percentage recalling longer stays is very similar to that found by Andersen *et al.* (1979) but the percentages claiming a shorter stay are less than they found. Those who agreed with the hospital records comprised 39.9% of the Andersen *et al.* study.

In our study, of those who overreported their lengths of stay, 61.1% erred by 1-2 days and 18.3% by six or more days. Of those who underreported their lengths of stay, 66.7%

overreported by 1-2 days and 21.8% by six or more days. These figures show the same trend but a wider range than Andersen *et al.* (1979).

The average reported length of stay was 10.9 days compared to 10.6 for the official records. This 0.3 difference equates to a 2.8% increase in the length of stay for the interview reported information. This is comparable to Simmons and Bryant's (1962) value of 2%.

The 'heaping effect' found by Cannell *et al.* (1965) for reported lengths of stay to be in multiples of 5 days was not found in our study, the peaks in our frequencies are at 7 days and multiples of this, and for physician data the largest peak is found at 12 visits per year. Our subjects seem to consider hospitalizations in weekly units, and physician visits in months.

Simmons and Bryant (1962) found that there were considerably more hospitalizations recorded for the most recent six months before the interview than in months 7-12. This is not the case with the present data where there are 152 visits in months 1-6 and 145 for months 7-12. There is no decline even as the 12th month is approached; there are as many visits in months 10, 11 and 12 as in the first three months.

As the months increase there is more underreporting, ranging from 2.5% in the first three months to 15.2% in

months 10-12. This is much less than Cannell *et al.* (1977) who found the percentage of underreports to range from 5% for less than two months to 50% near the 12th month.

For the hospitalizations, the overreporters were fairly reliable with the number of days reported and that recorded even though the hospitalizations the subject were recalling ranged from one to four years before the interview. Fifteen subjects who overreported hospitalizations were not hospitalized according to our records in the four years prior to the interview; of these 15, 11 reported stays of one day. These could be emergency or outpatient clinic visits reported as admissions in the interview. If overnight accommodation were provided in the emergency department this would not have been recorded as a hospital admission and therefore not included in the data obtained. For the overreporters, nine had visits in the 13th and 14th month before the interview; of these, six were for diagnoses of a chronic condition and therefore more likely to be prone to forward telescoping (Means and Loftus, 1991).

Comparing the over- with the underreporters, the overreporters are younger but there was no difference in sex, education or the health status variables. Cleary and Jette (1984) also found that younger people tend to report too many visits. On the other hand, those who agreed on having been hospitalized in the recall period are significantly younger,

more educated, more likely to be female and likely to suffer from fewer chronic conditions than the underreporters. This implies that the older, the males, the less educated and those with more chronic conditions underreport more than other subjects; these are similar findings relative to socio-demographic variables as previous studies (Cannell *et al.*, 1965; Cannell *et al.*, 1977; NCHS, 1965) with the exception that these authors found no effect of sex on underreporting. No significant differences are found between the overreporters and those in agreement on being hospitalized.

Within the agreement on hospitalization group, those who reported the length of stay accurately to within one day are more likely to be females, younger, with higher education, better self-assessed health status, fewer chronic conditions and better emotional health than those who show more discrepancy. This follows the premise that the older and less educated subject is less accurate in recall than others.

An accuracy variable was computed from a combination of 1) whether the subject was in agreement with official records, an under- or overreporter, together with 2) the degree of accuracy of reporting - including the number of months between the reported and recorded visit and the difference in the length of stay reported and recorded. This variable shows that the less accurate subjects are older, less educated, have lower self-assessed health status, a

higher number of chronic conditions and a lower emotional score. By sex, males show stronger associations for age, education and number of chronic conditions, and females for satisfaction with physical condition. When those hospitalized for pregnancy conditions are excluded the females approach the males in their associations for age, education and number of chronic conditions, and the association is stronger for self-assessed health status.

Logistic regression analysis re-emphasized the facts that 1) those who are more likely to disagree on hospitalizations having occurred are the less educated and those with a higher number of chronic conditions; 2) when the non-hospitalized are excluded from the analysis, males are more likely to disagree than females; 3) older subjects are more likely to underreport than the younger; and 4) for the accuracy variable, those with a higher number of chronic conditions, less education and a lower emotional score are more likely to be inaccurate.

Logistic regression in these analyses is not being used to predict outcome (as in questions such as "Will a subject cease smoking or not") and therefore odds and odds ratios are not applicable. Also, with very large numbers in one group of the dependent variable, the predictive powers of the program are not very robust and analysis of the groups

excluding the large group (e.g. excluding those not hospitalized in the recall period) may be more meaningful.

For physician visits, as would be expected, those who were hospitalized in the recall period made more visits to the physicians than those who were not hospitalized. Over half of all subjects made between zero and three visits in any year and almost three-quarters made between zero and six visits. As the number of visits increase so does the difference between the reported and recorded number, this may be due to 'generic memory' for events and the subjects simply getting more confused as the numbers of visits increase. Females make more visits than males; this is to be expected and has been shown many times in the literature. The difference is reduced if pregnancy related visits are excluded. The increase in the number of visits with increasing age and less education for males is not apparent for females, this could be because females visit physicians more in their younger years for consultations related to pregnancy which evens out with the known increase in visits with advancing age. Since age and education are correlated one would expect an increase in visits to be seen both in the older and less educated subjects.

Both overreporters and underreporters for physician visits are more likely to be male, have better self-assessed health status, fewer chronic conditions and have a better

emotional health than those in agreement; in addition, overreporters are more likely to be younger and have more education, and underreporters are more satisfied with their physical condition.

For accuracy of recalling the number of physician visits, 49% are accurate within one visit, but it must be remembered that over half the population make fewer than three visits in any year so it would be anticipated that they would remember reasonably accurately. The percentage overreporting is fairly constant irrespective of the number of visits made but there is an increase in underreporting and a corresponding decrease in those in agreement as the number of visits rises. Using a decomposition method (Jobe and Mingay, 1989) when a larger number of visits is reported may help with accuracy but its use would have to be weighed against the increase in the interview time and related costs.

Males are more accurate in recalling the number of visits than females, but they also make fewer visits and so have fewer errors since increasing number of visits is related to decreasing accuracy.

The computed accuracy variable for physician visits shows that there is decreasing accuracy in those older, less educated, lower self-assessed health status, a higher number of chronic conditions and a lower emotional score; age and chronic conditions are more associated in males than females.

These are the same circumstances as found for accuracy in hospitalization reporting.

Logistic regression for physician visits shows that increasing disagreement is related to higher self-assessed health status, fewer number of chronic conditions and being male. The underreporters have a lower education and fewer chronic conditions than the overreporters. Those in agreement have a lower self-assessed health status, higher education and more chronic conditions than the underreporters. As stated above, it is expected that those with a higher number of chronic conditions, lower self-assessed health status and a lower emotional score would be more likely to visit a physician and hence more likely to be in agreement on at least having had one visit. For accuracy of number of visits, being a female, having low self-assessed health status, lower education and a higher number of chronic conditions are all indicators of potential inaccuracy. This also follows the premise that all these states lead to an increase in the number of visits made and hence to more inaccuracy in the reported number.

Comparison between the accuracy variables for hospitalization and physicians shows a good association, particularly for males if those not hospitalized in the recall period are excluded. Whereas, the division of subjects into those in agreement, the over- and

underreporters shows no association between hospitalizations and physician visits.

The reassuring underlying result from this study is that most people report their hospitalizations and physician visits for the previous 12 months with a high degree of accuracy. Certain population groups are more inclined to inaccuracy and for these groups more detailed probing for the information might be beneficial. To increase the number of questions, and hence time taken for the interview, for all subjects in a general population survey would probably not be cost-effective.

The selection, training and supervision of interviewers is closely associated with response errors. Instructing interviewers to adhere strictly to the way questions are worded should help to achieve uniformity in the presentation and interviewer - respondent inter-reaction. Careful and accurate recording and coding of responses should be emphasized. Quality checks carried out by a supervisor should pick up possible error sources early in the survey.

REFERENCES

Andersen R, Kasper J, Frankel MR and Associates. (1979) Total Survey Error. Jossey-Bass, San Francisco.

Anderson OW. (1991) The evolution of health services research: personal reflections on applied social science. Jossey-Bass, San Francisco.

Baddeley A. (1979) The limitations of human memory: Implications for the design of retrospective surveys, in The recall method in social surveys, Moss L and Goldstein H (eds) University of London Institute of Education, London.

Belloc NB (1954) Validation of morbidity survey data by comparison with hospital records. J Am Stat Assoc December, 832-846.

BMDP. (1988) BMDP statistical software manual. BMDP Statistical Software, Inc. Los Angeles, USA.

Brewer MB, Dull VT and Jobe JB. (1989) Social cognition approach to reporting chronic conditions in health surveys. National Center for Health Statistics, Vital Health Stat 6(3).

Canada Health Survey. (1981) The Health of Canadians. Health and Welfare and Statistics Canada, Ottawa.

Cannell CF, Fisher G and Bakker T. (1965) Reporting of hospitalization in the Health Interview Survey. National Center for Health Statistics, Vital Health Stat 2(6).

Cannell CF, Marquis KH and Laurent A. (1977) A summary of studies of interviewing methodology. National Center for Health Statistics, Vital Health Stat 2(69).

Cartwright A. (1963) Memory errors in a morbidity survey. Milbank Q 41, 5-24.

Cherry N and Rodgers B. (1979) Using a longitudinal study to assess the quality of retrospective data, in The recall method in social surveys, Moss L and Goldstein H (eds), University of London Institute of Education, London.

Cleary PD and Jette AM. (1984) The validity of self-reported physician utilization measures. Med Care 22, 796-803.

Cohen B, Erickson P and Powell A. (1983) The impact of length of recall period on the estimation of health events. ASA Pro So St, 497-501.

Coulter A, McPherson K, Elliott S and Whiting B. (1985) Accuracy of recall of surgical histories: a comparison of postal survey data and general practice records. Community Med 7, 186-189.

Enterline PE and Capt KG. (1959) A validation of information provided by household respondents in health surveys. Am J Public Health 49, 205-212.

Fisher G. (1962) A discriminant analysis of reporting errors in health interviews, Appl Stat 11, 148-163.

Green S, Kaufert J, Corkhill R, Creese A and Dunt D. (1979) The collection of service utilization data: a research note on validity. Soc Sci Med 13(A), 231-234.

Harlow SD and Linet MS. (1989) Agreement between questionnaire data and medical records: The evidence for accuracy of recall. Am J Epidemiol 129(2), 233-248.

Jobe JB and Mingay DJ. (1989) Cognitive research improves questionnaires. Am J Public Health 79(8), 1053-1055.

Jobe JB, White AA, Kelley CL, Mingay DJ and Sanchez MJ. (1990) Recall strategies and memory for health-care visits. Milbank Q 68(2), 171-189.

Kars-Marshall C, Spronk-Boon YW and Pollemans MC. (1988) National Health Interview Surveys for Health Care Policy. Soc Sci Med 26(2), 223-233.

Kulley AM. (1974) The validity of survey measurement of health services utilization: a verification study of respondent reports of hospitalization. Ph.D. thesis, Purdue University.

Landis JR and Koch GG. (1977) The measurement of observer agreement for categorical data. Biometrics 33, 159-174.

Loftus EF. (1982) Memory and its distortions, in The G Stanley Hall lecture series, Volume 2, Kraut AG (ed). American Psychological Association, Washington, DC.

Manga P, Broyles RW and Angus DE. (1987) The determinants of hospital utilization under a universal public insurance program in Canada. Med Care 25(7), 658-670.

Mathiowetz NA and Groves RM. (1985) The effects of respondent rules on health survey reports. Am J Public Health 75, 639-644.

Means B, Nigam A, Zarrow M, Loftus EF and Donaldson MS. (1989) Autobiographic memory for health-related events. National Center for Health Statistics, Vital Health Stat 6(2).

Means B and Loftus EF. (1991) When personal history repeats itself: decomposing memories for recurring events. Appl Cognitive Psychol 5(4), 297-318.

Mechanic D. and Newton M. (1965) Some problems in the analysis of morbidity data. J Chronic Dis 18, 569-580.

Moser CA. and Kalton G. (1972) Survey methods in social investigation. Basic Books, Inc., New York.

NCHS. (1965) Comparison of hospitalization reporting in three health survey procedures. National Center for Health Statistics, Vital Health Stat 2(8).

Roghmann KJ and Haggerty RJ. (1974) Measuring the use of health services by household interviews: a comparison of procedures used in three child health surveys. Int J Epidemiol 3, 71-81.

Schach E, Kovacic L, Bice TW, Matthews VL, Haythorne DF, Paganini JM and Rabin DL. (1972) Methodologic results, in International comparisons of medical care, Rabin (ed) Milbank Q 50(3) Part 2, 65-80.

Segovia J, Bartlett RF, Veitch B and Edwards AC. (1987) Lifestyle, health practices and utilization of health services - Final Report. Memorial University of Newfoundland, St. John's, Canada.

Segovia J, Bartlett RF and Edwards AC. (1989) An empirical analysis of the dimensions of health status measures. Soc Sci Med 29(6), 761-768.

Segovia J, Bartlett RF and Edwards AC. (1991) Health Status and Health Practices - Alameda and beyond. Int J Epidemiol 20(1), 259-263.

Simmons WR and Bryant EE. (1962) An evaluation of hospitalization data from the Health Interview Survey. Am J Public Health 52, 1638-1647.

Statistics Canada. (1987) Census Canada 1986, Profiles, Newfoundland Part I. Census divisions and subdivisions. Cat. No. 94-101, Ottawa.

Sudman S and Bradburn NM. (1974) Response effects on surveys. Aldine Publishing Company, Chicago.

Veitch BM. (1991) A categorical data analysis of health practices, health status, and hospital utilization in metropolitan St. John's. MSc thesis, Memorial University of Newfoundland, St. John's.

WHO (1988). Manual of the international classification of diseases, 9th revision, clinical modification. World Health Organization, Geneva, Switzerland.

Wilson RW and Elinson J. (1981) National survey of personal health practices and consequences: background, conceptual issues and selected findings. Public Health Rep 96, 218.

Yaffe R, Shapiro S, Fuchsberg RR, Rohde CA and Corpeno HC. (1978) Medical economics survey-methods study - cost-effectiveness of alternative survey strategies. Med Care 16, 641-659.

Zaremba MM, Willhoite B and Ra K. (1985) Self-reported data: Reliability and role in determining program effectiveness. Diabetes Care 8, 486-490.

APPENDIX

INTERVIEWER ☐

HOUSEHOLD ☐☐☐☐

SUBJECT ☐☐☐

DATE OF INTERVIEW ☐☐☐☐☐☐

DATE RECEIVED ☐☐☐☐☐☐

BATCH NO. ☐☐

Memorial University of Newfoundland
Faculty of Medicine
DIVISION OF COMMUNITY MEDICINE AND
BEHAVIOURAL SCIENCES

LIFESTYLE, HEALTH PRACTICES AND MEDICAL CARE UTILIZATION SURVEY QUESTIONNAIRE

** TO BE REMOVED BY FIELD OFFICE BEFORE DATA PROCESSING

SUBJECT'S NAME_____TELEPHONE NO._____

ADDRESS_____

CONSENT STATEMENT READ ☐

Int. ☐ 1

I.D. Household ☐☐☐☐ 2

Subject ☐☐☐ 3

(M.C.P. NUMBER) FROM Q. 65 ☐☐☐ ☐☐☐ ☐☐☐ ☐☐☐ 4

WE ARE GOING TO ASK ABOUT YOUR EATING AND SLEEPING HABITS NOW.

1. How often do you eat breakfast?
   PROBE: *(EATING BREAKFAST MEANS MORE THAN A CUP OF COFFEE ONLY)*

   Every day, or almost every day 1 ☐
   Sometimes (3-4 times a week) 2 ☐
   Rarely, or never 3 ☐ 5

2. Do you make any conscious effort to limit the amount of red meat in your diet for *health reasons?*

   Yes 1 ☐
   No 2 ☐ 6

3. Do you make any conscious effort to limit the amount of animal fat in your diet?
   PROBE: (REAL BUTTER, WHOLE MILK, EGGS)

   Yes 1 ☐
   No 2 ☐ 7

4. How many hours do you sleep per night?
   (PROBE: SIX, SEVEN?) *CODE NUMBER OF HOURS DIRECT*

   ☐☐ 8

5. How tall are you?
   (WRITE THE ANSWER IN THE UNITS GIVEN BY THE RESPONDENT)

   ☐ feet ☐☐ • ☐ ins cm ☐☐☐ 9

6. How much do you weigh?
   RECORD THE ANSWER IN THE UNITS GIVEN BY RESPONDENT

   ☐☐☐ lbs kg ☐☐☐ 10

7. Do you consider yourself to be . . . .

   . . . overweight 1 ☐
   . . . underweight 2 ☐
   . . . about average 3 ☐ 11
   . . . DK 9 ☐

8. You are . . . Male 1 ☐ Female 2 ☐ 12

THE NEXT QUESTIONS ARE ABOUT SMOKING AND DRINKING

9. Did you ever smoke regularly?

PROBE. REGULAR SMOKING MEANS ONE CIGARETTE, PIPE, CIGAR A DAY FOR ONE YEAR

YES 1 ☐ NO 2 ☐ *GO TO Q 18* → 13

10. Are you smoking now?

YES 1 ☐ NO 2 ☐ *GO TO Q. 12* → 14

11. Do you smoke . . . . . .

	CURRENT Yes	CURRENT No	
. . . . . cigarettes	1 ☐	2 ☐	15
. . . . . pipe	1 ☐	2 ☐	16
. . . . . cigar	1 ☐	2 ☐	17

12. Did you ever regularly smoke . . . *ASK WHATEVER NOT MENTIONED ABOVE*

	EX	NEVER	
. . . . . cigarettes	1 ☐	2 ☐	18
. . . . . pipe	1 ☐	2 ☐	19
. . . . . cigar	1 ☐	2 ☐	20

IF YES TO ANY OF THE ABOVE

13 How old were you when you stopped smoking? *THE ABOVE*

. . . . . cigarettes	☐☐	21	
. . . . . pipe	☐☐	22	*CODE AGE DIRECT* NA 99
. . . . . cigar	☐☐	23	

ASK ALL SMOKERS PAST AND PRESENT

14. How old were you when you started smoking? *ASK WHATEVER MENTIONED*

. . . . . cigarettes	☐☐	24	
. . . . . pipe	☐☐	25	*CODE AGE DIRECT* NA 99
. . . . . cigar	☐☐	26	

15. During the period when you smoked most, how many cig/pipes cigars did you smoke a day?

. . . . . cigarettes	☐☐	27	
. . . . . pipe	☐☐	28	*CODE AGE DIRECT* NA 99
. . . . . cigars	☐☐	29	

16. Do/Did you inhale the smoke?

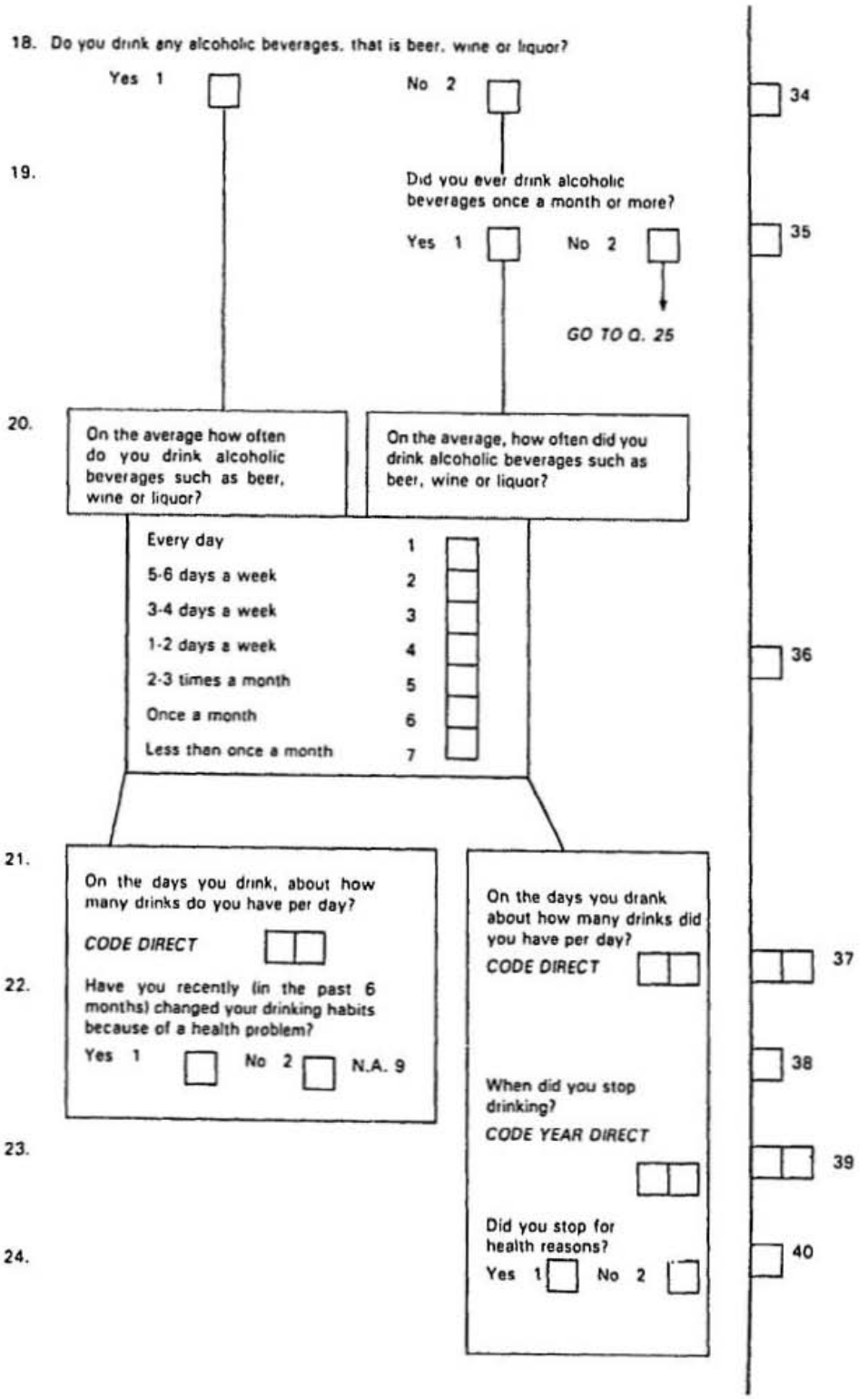
	Yes	No	
. . . . . cigarettes	1 ☐	2 ☐	30
. . . . . pipe	1 ☐	2 ☐	31
. . . . . cigar	1 ☐	2 ☐	32

NA 99

ASK CURRENT SMOKERS

17. During the past two years, did you make a serious attempt to stop smoking?

Yes 1 ☐ No 2 ☐ NA 9 → 33


18. Do you drink any alcoholic beverages, that is beer, wine or liquor?
Yes 1
No 2
34
19.
Did you ever drink alcoholic beverages once a month or more?
Yes 1
No 2
35
GO TO Q. 25
20.
On the average how often do you drink alcoholic beverages such as beer, wine or liquor?
On the average, how often did you drink alcoholic beverages such as beer, wine or liquor?
Every day 1
5-6 days a week 2
3-4 days a week 3
1-2 days a week 4
2-3 times a month 5
Once a month 6
Less than once a month 7
36
21.
On the days you drink, about how many drinks do you have per day?
CODE DIRECT
22.
Have you recently (in the past 6 months) changed your drinking habits because of a health problem?
Yes 1
No 2
N.A. 9
On the days you drank about how many drinks did you have per day?
CODE DIRECT
37
38
When did you stop drinking?
CODE YEAR DIRECT
23.
39
Did you stop for health reasons?
Yes 1
No 2
24.
40

THE NEXT SECTION IS ABOUT YOUR PHYSICAL ACTIVITIES

25. Are you now suffering from any disability

(PROBE: A CONDITION THAT STOPS YOU FROM DOING YOUR ROUTINE ACTIVITIES)

Yes 1 ☐
No 2 ☐ *GO TO Q 29* ☐ 41

26. Is it a temporary condition?

(PROBE: A CONDITION THAT WILL DISAPPEAR IN A FEW WEEKS)

Yes 1 ☐
No 2 ☐ ☐ 42
DK 9 ☐

27. Was it caused by an accident or injury?

Yes 1 ☐
No 2 ☐ *GO TO Q 29* ☐ 43

28. Did this accident or injury happen . . . .

. . . . at home 1 ☐
. . . . outdoors 2 ☐
. . . traffic 3 ☐ ☐ 44
. . . . at work 4 ☐

29. How many times in a 2 week period do you usually do any of the following exercises or recreational activities?

How much time did you spend on each occasion?

READ	No. of Times	Mins 1-10	Mins 15+	NA				
1. Walking (including to and from school or work)	☐☐	1	2	9	45	☐☐	46	☐
2. Jogging or running	☐☐	1	2	9	47	☐☐	48	☐
3. Calisthenics (doing physical exercises)	☐☐	1	2	9	49	☐☐	50	☐
4. Bicycling (including to and from work)	☐☐	1	2	9	51	☐☐	52	☐
5. Bowling	☐☐	1	2	9	53	☐☐	54	☐
6. Vigorous dancing	☐☐	1	2	9	55	☐☐	56	☐
7. Skating	☐☐	1	2	9	57	☐☐	58	☐
8. Team sports (such as baseball, softball etc)	☐☐	1	2	9	59	☐☐	60	☐
9. Swimming	☐☐	1	2	9	61	☐☐	62	☐
10. Gardening	☐☐	1	2	9	63	☐☐	64	☐
11. Racquet sports	☐☐	1	2	9	65	☐☐	66	☐
12. Golf	☐☐	1	2	9	67	☐☐	68	☐
13. Other (Specify _____)	☐☐	1	2	9	69	☐☐	70	☐

30 Are you more, less, or equally active in winter?

More 1 ☐
Less 2 ☐
Equally 3 ☐ 71

AND NOW SOME QUESTIONS IN RELATION TO MEDICAL CARE

31 In the last year, that is from of 1984, did you have a consultation with a doctor?

Yes 1 ☐
No 2 ☐ *Go to Q 33* 72

32. How many visits did you have in the last year?

CODE DIRECT
NA or DK 99 ☐☐ 73

33. Do you have a family doctor?

(PROBE: A DOCTOR WHOM YOU ALWAYS CONSULT)

Yes 1 ☐
No 2 ☐ 74

34. Within the last year (from 1984) have you stayed at home because of an illness, or not feeling well

Yes 1 ☐
No 2 ☐ *GO TO Q. 37* 75

35. Did you stay in bed?

Yes 1 ☐
No 2 ☐ 76

36. How many days did you stay in bed?

CODE DIRECT ☐☐☐ 77

37. In the last year (same period) have you been a patient in a hospital overnight?

(PROBE: DID YOU SPEND AT LEAST A NIGHT IN A HOSPITAL)

Yes 1 ☐
No 2 ☐ *GO TO Q. 40* 78

38. How many days did you spend at the hospital?

CODE DIRECT ☐☐☐ 79

FOR FEMALES ONLY

39. Was the hospitalization due to pregnancy or delivery?

Yes 1 ☐
No 2 ☐ 80

40. We would like to know how satisfied or dissatisfied you are, in general with medical care in your own experience. On a five-point scale in which 5 means that you are very satisfied, and 1 means that you are very dissatisfied, what will be your score?

(PROBE: THINK OF A LADDER WITH FIVE RUNGS, WHERE THE HIGHEST OF THE FIVE IS THE BEST, WHERE ARE YOU ON THIS LADDER?)

SATISFIED

5
4
TICK 3
2
1

DISSATISFIED

DK 9 ☐ ☐ 81

NOW LETS GO BACK TO YOUR OWN HEALTH AND WELL BEING

41. Would you say that your yealth is . . . .

. . . . Excellent	1	☐
. . . . Good	2	☐
. . . . Fair	3	☐
. . . . Poor	4	☐

☐ 82

42. Over the past year, has your health caused you. . . .

. . . . no worry at all	1	☐
. . . . Hardly any worry	2	☐
. . . . Some worry	3	☐
. . . . A great deal of worry	4	☐

☐ 83

43. Do you have any of the following chronic conditions?

(CHRONIC MEANS THE CONDITION HAS BEEN PRESENT FOR THREE MONTHS OR MORE)

READ LIST;

CIRCLE CODES THAT CORRESPOND

Anemia	01	High blood pressure	13
Allergy (OF ANY KIND)	02	Kidney disease (stones etc.)	14
Arthritis, rheumatism	03	Mental illness	15
Asthma	04	Missing arm(s) or leg(s)	16
Cancer	05	Missing finger(s) toes	17
Cerebral Palsy	06	Paralysis of any kind	18
Diabetes	07	MALES: Prostrate disease	19
FEMALES: Dysmenorrhea (menstrual problems)	08	Recurring backaches	20
		Recurring headaches	21
Emphysema	09	Stomach ulcer	22
Epilepsy	10	Thyroid trouble or goitre	23
Heart disease	11	Tuberculosis	24
Hemorrhoids (piles)	12	OTHER	
		Specify_____	25
		None	88

☐☐ 84
☐☐ 85
☐☐ 86
☐☐ 87
☐☐ 88
☐☐ 89
☐☐ 90
☐☐ 91

44. Compared with other people your age, would you say you have . . . . .

. . . . much more energy	1	☐
. . . . somewhat more (energy)	2	☐
. . . . average amount of energy	3	☐
. . . . somewhat less (energy)	4	☐
. . . . much less energy	5	☐

☐ 92

45. In general, how satisfied are you with your overall physical condition. . . .

. . . . are you very satisfied	1	☐
. . . . satisfied	2	☐
. . . . not too satisfied	3	☐
. . . . not at all satisfied	4	☐

☐ 93

46. During the past few weeks, how often have you felt. . . .
CIRCLE

would you say	Often	Sometimes	Never	
. . . . on top of the world	1	2	3	☐ 94
. . . . lonely	1	2	3	☐ 95
. . . . that things were going your way	1	2	3	☐ 96
. . . . restless	1	2	3	☐ 97
. . . . depressed, or unhappy	1	2	3	☐ 98

47. All in all, how happy are you these days? Would you say. . .

. . . . very happy	1	☐
. . . . pretty happy	2	☐
. . . . not too happy	3	☐
. . . . unhappy	4	☐

☐ 99

48. How many close relative do you have? These are people that you feel at ease with, can talk to about private matters, and can call on for help. *(DO NOT INCLUDE SPOUSE)*

CODE DIRECT

☐☐ ☐☐ 100

49. How many close friends do you have? These are people that you feel at ease with, can talk to about private matters and can call on for help.

CODE DIRECT

☐☐ ☐☐ 101

NOW SOME QUESTIONS ABOUT PREVENTIVE HEALTH:

50. How often do you brush your teeth?
PROBE (DO YOU HAVE YOUR OWN TEETH?)

PROBE: More than twice a day 1 ☐
Twice a day 2 ☐
Once a day 3 ☐
Not every day 4 ☐
No teeth Code NA 9 9 ☐ *GO TO Q 54* → ☐ 102

51. Do you use dental floss? (WATER PICK COUNTS AS FLOSS)

Yes 1 ☐
No 2 ☐ *GO TO Q 53* → ☐ 103

52. How often?

Every day 1 ☐
Every week 2 ☐ ☐ 104

53. When was the last time that you went to a dentist?

USE LIST AS PROBE:
Within the last year 1 ☐
one to two years 2 ☐
more than two years 3 ☐
Never 4 ☐
DK 9 ☐ ☐ 105

54. When was the last time that your went to a doctor for a preventive examination when you were not sick?

PROBE: FEMALES: PAP SMEAR, BREAST EXAMINATION
MALES: BLOOD PRESSURE CHECK

CODE YEAR DIRECT

☐☐ ☐☐ 106

Never 00
DK 99

55. Do you use your seatbelt while travelling by car?

Yes 1 ☐
No 2 ☐ ☐ 107

TO COMPLETE THE QUESTIONNAIRE WE NEED A FEW MORE DETAILS:

56. Where were you born? Was it Newfoundland?

IF CANADA, ASK PROVINCE IF NOT IN CANADA ASK COUNTRY. CIRCLE

NFLD.	01	MAN.	07	U.K.	13
N S.	02	SASK.	08	U.S.A.	14
N.B.	03	ALBTA	09	OTHER	
P.E.I.	04	B C.	10	AMERICAS	15
QUE.	05	YUKON	11	EUROPE	16
ONTARIO	06	N.W.T.	12	ASIA	17
				OTHER	18

☐☐ 108

57. What is your marital status?
PROBE: ARE YOU MARRIED?

Single	1 ☐
Married	2 ☐
Divorced/Separated	3 ☐
Widowed	4 ☐

☐ 109

58. What was the last grade you complete in school?

CODE DIRECT

☐☐

☐☐ 110

59. ASK ONLY IF ANSWER INDICATES THAT RESPONDENT COMPLETED HIGH SCHOOL

Do you have any education beyond High School?

Yes	1 ☐	
No	2 ☐	*GO TO Q. 62*

☐ 111

60 What kind of education was it?

Tradeschool, diploma courses etc.	1 ☐	*GO TO Q. 62*
University	2 ☐	

☐ 112

61. Do you have a university degree?

Yes	1 ☐	
No	2 ☐	

☐ 113

62. Are you now.... working 1
... retired 2
CIRCLE .... unemployed 3
.... laid off temp/on strike 4
.... unable to work (disability) 5
... keeping house 6
... studying 7

GO TO Q. 64

☐ 114

63. What is/was your job?
PROBE: WHAT DO YOU DO AT WORK?

_____ 115

64 What is your date of birth?

Y Y M M D D 116

65. What is your M C.P. No ?

ASK HUSBAND AND WIFE ONLY

66. What is the approximate total income for your household?
(PROBE INCLUDING ALL WAGES, SALARIES, PENSIONS, AND ALLOWANCES)

. . . . is it less than $15,000 1
. . . . between $15,000 and $30,000 2
. . more than $30,000 3 119
N.A. 9

THAT COMPLETES THE INTERVIEW. THANK YOU VERY MUCH FOR DONATING YOUR TIME TO THE STUDY. IT IS VERY MUCH APPRECIATED.

TO BE CODED FOLLOWING COMPLETION OF INTERVIEW, FROM THE HOUSEHOLD SHEET

CODE EITHER TO WIFE (OR SINGLE FEMALE) OR HUSBAND (OR SINGLE MALE)

CODE 9 for all the rest – DO NOT LEAVE BLANKS

	Deceased 1	Independent 2	Family 3	Nursing H 4	H H 5	
67. WIFE'S Mother						120
Father						121
68. HUSBAND'S Mother						122
Father						123

CODE FROM HOUSEHOLD SHEET, AFTER COMPLETION OF ALL INTERVIEWS

69. Total number of subjects in H.H. 124
Total number of children 19 or less 125

TO BE CODED ONLY BY FIELD OFFICE:

70. Total number of refusals 126
Total number of non-respondents 127

Table A1
Frequency distributions of socio-demographic variables

	n	%
SEX		
male	1349	45.1
female	1645	54.9
AGE (grouped)		
20 - 44	1992	66.5
45 - 64	682	22.8
65+	320	10.7
EDUCATION (grouped)		
less than high school	899	30.0
high school	426	14.2
college/diploma	820	27.4
university - no degree	401	13.4
university - with degree	442	14.8
missing information	6	0.2

Table A2
Frequency distributions of health status variables

	n	%
SAHS - self-assessed health status		
excellent	824	27.5
good	1616	54.0
fair	509	17.0
poor	45	1.5
CHRCOND - number of chronic conditions		
none	1291	43.1
1	855	28.6
2	482	16.1
3	212	7.1
4	103	3.4
5	31	1.0
6	14	0.5
7	5	0.2
8	1	0.0
PHYSCOND - satisfaction with physical condition		
very satisfied	544	18.2
satisfied	1901	63.5
not too satisfied	496	16.6
not at all satisfied	52	1.7
missing information	1	0.0
EMOT - emotional status		
excellent	1381	46.1
good	1316	44.0
fair	248	8.3
poor	39	1.3
missing information	10	0.3

Table A3
Hospitalizations - listing of overreports where subject was not hospitalized at any time in the 5 years previous to the interview

ID	Age (yrs)	Sex	Educ (yrs)	Days reported in hospital
161102	59.9	M	7	1
304101	34.2	F	13	3
450102	36.1	M	10	14
525101	63.4	F	7	1
645101	70.5	F	10	1
729102	40.2	M	13	1
788101	41.7	F	4	7
803501	29.0	F	12	2
862205	23.7	F	12	1
920101	42.8	F	11	1
965203	25.0	M	12	1
978102	41.6	M	13	1
1112102	25.6	M	15	1
1475403	25.7	F	13	1
1674601	49.7	M	12	1

Table A4
Hospitalizations - listing of underreports

ID	Age (yrs)	Sex	Educ (yrs)	No. days before interview	Length of stay	Diagnosis
155102	65.1	M	11	-355	11	Atherosclerosis
192101	70.4	F	3	-313	19	Intestinal obstruction
202102	52.6	M	5	-346	28	Alcohol dependency syndrome
328403	24.5	F	15	-325	12	Gall bladder
491202	25.9	M	12	-395	2	Cardiac dysrhythmia
537102	34.1	M	13	-154	3	Vasoplasty
564101	66.7	F	11	-292	2	Injury - face/neck
574102	52.9	M	9	-338	4	Bronchitis
620102	66.2	M	5	-200	2	Genitourinary - phimosis
730102	66.7	M	11	-210	3	Chronic airway obstruction
"				-277	12	Abdominal aneurysm
"				-306	9	Heart failure
798101	76.2	F	6	-342	12	Gastric ulcer
916601	69.3	F	6	-349	6	Swelling in head/neck
943101	33.6	F	15	-63	1	Excessive menstruation
"				-297	4	Normal delivery
1139102	69.1	M	8	-144	8	Fracture of ankle
1284101	25.4	F	10	-389	4	Bronchitis
1397101	25.3	F	12	-276	4	Normal delivery
1436102	80.8	M	4	-300	20	Gall bladder
1492601	69.3	F	8	-239	16	Genitourinary - fistula
"				-350	14	Colostomy
1595102	62.7	M	11	-276	5	Inguinal hernia
1738203	21.0	M	12	-82	7	Ulcer of lower limb
1854303	72.0	F	8	-292	6	Disease of salivary glands
1961102	70.5	M	11	-254	3	Senile cataract
2041102	50.4	M	13	-131	21	Alcohol dependency syndrome

Table A5
Hospitalizations - listing of overreports where subject was hospitalized sometime in the 5 years previous to the interview
(where more than one visit, most recent one listed)

ID	Age (yrs)	Sex	Educ (yrs)	Days difference	Length of stay hosp	interview	Diagnosis
4101	31.3	F	10	-409	4	3	Normal delivery
90102	25.2	M	10	-1152	2	2	(not available)
219101	62.5	F	10	-414	2	3	Hernia
244101	73.5	F	9	-399	4	7	Diabetes
303101	25.1	F	9	-443	25	30	Pre-eclampsia
411101	68.0	F	6	-908	21	33	Gallbladder
481102	63.9	M	8	-1207	16	4	Fracture of ankle
526601	38.1	M	5	-769	12	30	Orchitis
622205	22.4	F	8	-1471	5	3	Infl. dis pelvic organs
671303	69.9	M	8	-500	116	120	Acq. deformity of toe
693102	34.2	M	13	-911	8	1	Frac. of tibia/fibula
723101	30.3	F	12	-825	4	1	Ectopic pregnancy
808102	46.6	M	12	-655	3	3	Salmonella
818203	20.7	M	10	-379	6	4	Regional enteritis
872601	73.6	F	11	-381	9	20 *	Phlebitis
993203	31.3	M	6	-443	2	3	Ureteric stone
993205	22.3	M	6	-749	75	75	Intercranial hemorrhage
1011101	68.4	F	6	-585	2	1	Dysphagia
1019101	64.2	F	3	-393	21	21	Chr. isch. heart dis.
1034502	21.5	F	13	-381	7	7	Ovarian cyst
1061101	58.7	F	5	-477	12	14	Chr. isch. heart dis.

* 2 close hospitalizations, readmitted after 5 days; total days 22+9 = 31

Table A5 - continued
Hospitalizations - listing of overreports where subject was hospitalized sometime in the 5 years previous to the interview
(visit closest to interview)

ID	Age (yrs)	Sex	Educ (yrs)	Days difference	Length of stay hosp	interview	Diagnosis
1191202	24.0	F	11	-606	13	22	Disproportion (preg)
1245102	53.6	M	6	-411	32	28	Atherosclerosis
1262601	71.2	M	9	-585	4	5	Diabetes
1275502	20.5	M	13	-639	1	2	Aortic valve disorders
1333204	20.7	M	8	-504	4	2	Inguinal hernia
1510101	26.9	F	11	-424	4	4	Prem. rupt. membranes
1519201	20.9	F	13	-620	2	1	Asthma
1578101	33.7	F	11	-947	4	5	Tonsillitis
1581102	25.7	M	12	-436	6	7	Anomaly of jaw
1596101	34.2	F	12	-1068	4	1	Abdominal pain
1753102	31.2	M	11	-826	7	7	Pneumonia
1820601	87.2	F	11	-602	5	4	Hemorrhoids
1839102	37.4	M	10	-730	2	2	(not available)
1880101	36.9	F	12	-615	5	7	Dysmenorrhea
1925101	52.6	F	12	-566	2	2	Abdominal pain
1941101	49.3	F	11	-1062	5	16	Ovarian cyst
1959601	76.8	F	12	-777	3	4	Polyp on uterus
1977101	35.0	F	12	-940	6	2	Diabetes comp. preg.
2024101	28.2	F	12	-611	9	7	Ureteric stone
2044101	35.0	F	15	-917	8	1	Uter. scar -prev. surg.
2081501	21.3	F	13	-624	3	2	Tonsillitis

Table A6
Hospitalizations - underreports,
health status variables

	All under-reporters	Excluding vasoplasty and alcohol cases	Excluding all sensitive diagnoses
SAHS			
$\bar{x}$	2.2	2.2	2.2
SD	1.0	1.0	0.9
SE	0.2	0.2	0.2
(n)	(23)	(20)	(18)
CHRCOND			
$\bar{x}$	2.2	2.2	2.2
SD	1.7	1.6	1.6
SE	0.4	0.4	0.4
(n)	(23)	(20)	(18)
PHYSCOND			
$\bar{x}$	2.1	2.1	2.1
SD	0.7	0.7	0.7
SE	0.1	0.2	0.2
(n)	(23)	(20)	(18)
EMOT			
$\bar{x}$	2.0	1.9	1.9
SD	0.8	0.9	0.9
SE	0.2	0.2	0.2
(n)	(21)	(18)	(17)

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

Table A7
Hospitalizations - overreports,
health status variables

	All over-reporters	Those in hospital since 1981	Those never in hospital since 1981
SAHS			
$\bar{x}$	2.1	2.1	2.1
SD	0.7	0.8	0.7
SE	0.1	0.1	0.2
CHRCOND			
$\bar{x}$	1.6	1.5	1.7
SD	1.6	1.5	1.9
SE	0.2	0.2	0.5
PHYSCOND			
$\bar{x}$	2.0	2.0	2.1
SD	0.7	0.7	0.7
SE	0.1	0.1	0.2
EMOT			
$\bar{x}$	1.8	1.9	1.7
SD	0.7	0.7	0.7
SE	0.1	0.1	0.2
(n)	(57)	(42)	(15)

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

Table A8
Hospitalizations - overreports where subjects were hospitalized sometime prior to year before interview, health status variables

	All over-reporters	Those in 13th or 14th month	Those in 15th or greater month
SAHS			
$\bar{x}$	2.1	2.4	2.1
SD	0.8	0.5	0.8
SE	0.1	0.2	0.1
CHRCOND			
$\bar{x}$	1.5	1.8	1.5
SD	1.5	1.6	1.5
SE	0.2	0.5	0.3
PHYSCOND			
$\bar{x}$	2.0	2.1	2.0
SD	0.7	0.6	0.8
SE	0.1	0.2	0.1
EMOT			
$\bar{x}$	1.9	2.1	1.8
SD	0.7	0.6	0.7
SE	0.1	0.2	0.1
(n)	(42)	(9)	(33)

SAHS: self-assessed health status
CHRCOND: number of chronic conditions
PHYSCOND: satisfaction with physical condition
EMOT: emotional status

Table A9
Physician visits - frequency distribution for official record of physician visits (N=2994)

Number of visits	Number of subjects	Number of visits	Number of subjects
0	560	22	7
1	417	23	2
2	352	24	5
3	314	25	3
4	225	26	5
5	189	27	4
6	173	28	1
7	106	29	1
8	117	30	2
9	86	31	2
10	83	32	3
11	62	33	1
12	45	34	1
13	58	36	1
14	41	37	2
15	17	38	1
16	25	39	1
17	22	41	1
18	22	44	1
19	15	47	1
20	12	70	1
21	6	104	1

Table A10
Physician visits - percentage of subjects at various levels of discrepancy for different levels of recorded visits

	NUMBER OF RECORDED VISITS				
	0 - 3	4 - 6	7 - 15	16+	Total
ACCURACY OF REPORTING					
Underreport ≥1 vis	24.9	65.4	69.9	77.6	44.7
Agreement	36.9	10.1	4.9	2.0	23.4
Overreport ≥1 vis	38.1	24.6	25.2	20.4	31.9